UiO : Department of Informatics
University of Oslo

Crowdpinion

Obtaining people's momentary opinion in dynamic environments

Master's thesis

Marek F. Machnik

Master's Thesis

Informatics: Programming and networks

Department of Informatics, University of Oslo

Title:

Crowdpinion - Obtaining people's momentary opinion in dynamic environments

Author:

Marek Franciszek Machnik

Supervisors:

Sagar Sen (Simula Research Laboratory)

Magne Jørgensen (Simula Research Laboratory, University of Oslo)

Oslo, June 2015

Contents

1. Introduction

Many interesting results can be obtained for analysis in studies of momentary opinions, feelings, emotions – *the things that occur in human mind for a brief moment and are often gone shortly afterwards*. This is valid for studies from various areas such as Informatics [1] [2], Medicine [3] [4], Sociology, Psychology [5] etc. and for market studies [6] as well.

The data from momentary self-reporting studies is often more valuable than retrospective overall data collected sometime after the studied period. For example, it is quite common in big businesses to ask the employees periodically to give an opinion about their working environment. The employees are then given a long survey sheet with many questions about how have they felt at work in last e.g. 6 months. This method has a couple of downsides. It is heavily based on participants' long term memory and is prone to being affected by some contemporary biases, e.g. if such survey is conducted at a stressful month before deadlines, the results are likely to be worse than if it was conducted shortly after summer holidays or company-wide salaries rise. In an experience sampling study, if the employees are asked to answer small sets of questions a couple of times during a day in a period of one or two months, the results of the study will give more reliable and unbiased results [5].

Long-term Experience Sampling Method (ESM) studies bring the benefits mentioned above, but at the same time create several challenges. In this thesis, I focus on two of them:

- **How to obtain momentary opinion from the users effectively?**
- **How to keep participants engaged in a long-term study?**

This thesis addresses these questions by performing a systematic literature review, developing a tool, Crowdpinion, to garner momentary opinion from people participating in long-term studies. I conduct several studies to evaluate it in the context.

I have gone through various topics and branches of the domain. I have started with a strictly psychology-oriented tool that was supposed to detect the state of *flow* by asking people about their feelings at random moments. Then the idea evolved and I oriented my work towards sampling the momentary experience in order to evaluate User Experience (UX) of some software products. At that stage, I have conducted a systematic literature review [7] on momentary UX evaluation methods, which gave me an overview of the state of research in this field. The most recent work has led to creation of Crowdpinion, an ESM surveying system with a quite wide range of applications. While creating the tool I have done some research in the fields of crowdsourcing and gamification. All the ideas and topics have come together to create this thesis.

1.1. Crowdpinion: a tool for obtaining momentary opinion

Since I started working on my thesis, even if the topic has been more or less fixed with only minor alterations, the system, which I have been developing, has been evolving quite remarkably, especially at the conceptual level. There were three main stages and I think it is important to briefly mention them all. The paragraphs below explain the evolution of the system in three main stages. The last paragraph gives more information about what Crowdpinion is now.

ESM tool for detecting flow

It all started with a book by Mihály Csíkszentmihályi [8] and an idea to create a mobile app that would implement the Experience Sampling Method to query people in different groups (office workers, members of a rowing club, patients in a hospital) about their emotions ("Are you excited?", "Are you sad?" etc.) in random moments. The results could be used to detect *flow* – a particular state of mind, described by Csíkszentmihályi, when a person is totally immersed in their activities and performs them very effectively without being distracted by external factors. At that stage, detection of the flow has been my key focus while designing the system.

ESM tool for Momentary User Experience evaluation

After some weeks, I have realized that this kind of tool can be used to evaluate User Experience of other products or more precisely software products [9]. Immersion is in fact one of the positive UX factors [10]. In the new approach, the ESM studies in the app would be designed in a way that the users would be asked about their emotions while using a software product. This means that the users would be using the evaluated software and their smartphones would 'beep' every now and then asking them to respond a couple of questions like "Do you find using this software enjoyable?" or "Does the software help you in doing your tasks?"

Universal event-contingent ESM tool for various studies

At some point of working on the momentary UX evaluation tool, I thought that technically nothing limits the tool to the UX evaluation domain. The questions that would be asked to the app users could be about any research topic, not only emotions or opinions about software. At this stage, an extended study administration panel has been added to the tool, making it easy to create studies about various topics. Then I have also switched to the event-contingent ESM protocol, finding it more optimal for a wide range of studies. Then the software became a universal event-contingent ESM surveying tool known as Crowdpinion and some gamification [11] elements have been added to increase the participants' engagement in long-term studies.

This lead to what Crowdpinion is today: a surveying system for studies on momentary opinion in a wide range of topics. The system consists of a mobile app used by the participants to respond to questions and a web panel for the researchers, where they can set up and control their studies. The app includes some elements of games (leaderboard, unlockable features) and by doing this, it applies the idea of gamification

to make the participants more engaged. Several examples of studies, in which the researchers can apply Crowdpinion, can be found in the next subchapter. Extended description of the features of Crowdpinion is in the chapter 3. Of this thesis.

1.2. Crowdpinion: Why do we need to garner momentary opinion?

I created Crowdpinion because I believe that the studies on momentary opinion, feelings and experience can add valuable input to the more common retrospective studies. The Experience Sampling studies can be combined with these classic studies or conducted as separate studies. Below I have described three possible topics of such studies. There are of course many other possible applications.

1.2.1. Case 1: Work environment

One of the potential applications of Crowdpinion is evaluation of attitude at work throughout the day. This kind of study can be conducted as a research project in order to obtain data for a paper about general feelings and opinions of employees at work. However, it can be also ordered by a company or organization in order to obtain the data about the employees' attitude and opinions about the daily affairs. This data can then be used in process improvement to identify and resolve problems pointed out by the employees.

As the study aims to analyze employees' emotions during various moments of a working day, the list of events can include:

- Travelling to work
- Arriving in the office
- Just before a meeting
- Just after a meeting
- Completing a task
- Coffee break
- Lunch break
- Chat with a colleague
- Mailing
- Getting ready to go home
- Travelling home

The participants shall be asked to open the app every time when one of the events from the list occurs, select the event from the list and answer the questions that will appear on the screen. In real world, there is a risk that the employees will not have time or will to answer the questions every time when an event occurs. However, if they react to most of the events, we think that the results will still be valuable.

A set of questions asked at one of the events, e.g. the "Just after a meeting" event can consist of following questions:

1. Do you think the meeting was productive?
   (Neg.: Waste of time; Pos: Very productive)
2. Are you motivated to work in the remaining hours?
   (Neg.: Demotivated; Pos: Motivated)
3. Are you sleepy?
   (Neg.: Sleepy; Pos.: Full of energy)

The two first questions clearly are a part of the study. They can provide the researcher with meaningful data about the employees' motivation and perception of the meetings. The results can be even more interesting if similar questions are asked e.g. before and after a meeting. If such study shows that employees' motivation tends to drop after meetings and the meetings are mostly evaluated as "waste of time", then the management should rethink the way the meetings are organized or even reduce the number of meetings.

The third question ("Are you sleepy?") is not a part of the study. It is one of the questions that are meant to provide some entertainment to the participants and attract their attention. While the researchers obtain valuable data about meetings' effectiveness, the employees may find it entertaining to express how sleepy they are after a meeting and check if the other employees feel the same. This question could have been added to the study by the researcher or by one of the participants as one of the gamification bonuses.

While responding to questions the participants gain access to summary of responses to the questions. For example, at the beginning they can see the responses to one question (e.g. "Do you think the meeting was productive? – Just after a meeting") and information about the number of responses required to unlock another summary (e.g. "Give 10 more responses and unlock summary of 'Are you satisfied with the effects of your todays work? – Travelling home'."). The issues included in the study are interesting for the employees, so they are motivated to respond questions in order to see new summaries.

I have included this use case description in my paper [12] written and published together with Michael Riegler and Sagar Sen at The Second International Workshop on Gamification for Information Retrieval (GamifIR'15)

1.2.2. Case 2: Services at Oslo S

Let us imagine that the managers of Oslo S (Oslo Central Station) is planning to spend some money on improving the services the station want to research passengers' opinion about the comfort of travelling and the quality of services in order to find the areas that could be improved. Crowdpinion can be used for such study as well.

In this case, the group of potential responders is much broader than the group of employees from the first case. Practically every traveler can be an active study participant if they are asked to install the app. Oslo S should therefore advertise the study – e.g. by distributing leaflets and putting posters on the platforms. The app shall be publicly available in the Play Store and App Store, so everyone can easily install it in their smartphone devices. If the study is well advertised, the organizers can gather thousands of participants without a big effort.

The participants would probably belong to one of these two groups: long term participants (people who travel to and from Oslo S regularly, e.g. while commuting to work) and short term participants (people who use trains so rarely that they would probably be at Oslo S just once or twice during the period of the study). The second group includes the tourists as well. They can join the study at any time, even weeks after it started. These two groups will use the app differently, as the short term participant will not appreciate the gamification elements, but for the study the responses from both groups should be equally valuable.

In this study the events list can reflect the steps of the typical "flow" of the passengers in the station. It can include:

- Entering the building
- Arriving at the main hall
- Buying train tickets
- Buying a coffee
- Checking the departures table
- Finding the way to the platform
- Waiting for the train
- Hearing an announcement regarding the train
- Boarding the train

A sample set of questions for the Checking the departures table event can be similar to the following:

1. Was it easy to find your train?
   (Neg. Nearly impossible; Pos: Very easy)
2. Is the table readable?
   (Neg. Hard to read; Pos: Readable)

3. Are you confused?
   (Neg. Confused; Pos: Not confused)
4. Would you have a coffee before heading to the train?
   (Neg: No; Pos: Yes)

The first two questions ask about usability of the departure information table. The third one allows the passenger to express their feelings, which in this case are also related to the usability. “Would you have a coffee before heading to the train?” is an example of a question that can be used to analyze the flow – if many passengers state that they buy coffee before boarding the train, perhaps coffee machines should be installed at the platforms?

It is easy to see that by using the tool the study organizers would remarkably reduce the cost and effort of organizing such study. An alternative would perhaps be to hire a team of people who would fill the questionnaires with the passengers on the platforms, which would be costly and not necessarily effective, considering the usual attitude towards street surveys.

1.2.3. Case 3: User Experience of a software product

As mentioned before, the software that is now Crowdpinion, was earlier aimed to be a momentary User Experience (UX) evaluation tool. Crowdpinion in its current, more universal state, is still well suited for this kind of studies. I believe that it can successfully be applied in software development to support the UX designers and analysts in their work with various software.

In this case, the participants are most likely a narrow group selected by the UX specialists. The study can be done in two main ways. It can be supervised and take place in a UX lab, an office or any room where the supervisors would have direct contact with the participants (alpha testing [13]). The participants would be using the evaluated software according to the supervisors’ instructions and answer the questions in the app at specific moments. In this type of study, the role of Crowdpinion would perhaps be limited to the tool for collecting data from the supervised study. It still would help the study organizers to gather the results in one database and analyze them in an easy way.

Crowdpinion can be applied more fully in the other type of UX tests – the beta tests [13]. This type of tests is by definition performed at the client’s site, e.g. at the client’s office or in the users’ homes. These studies are not supervised and can be conducted over a long time. The study organizers can ask the users to follow some specific scenarios or use the evaluated software in the casual way. Crowdpinion would be the core of such study, being the main mean of communication between the UX specialists and the evaluators. The study organizers would have continuous control over the users testing the software in their homes or offices.

In both alpha and beta studies the configuration of Crowdpinion would be similar. The events list should cover the moments during the evaluated software’s workflow that are

particularly interesting for the UX experts. An example list for a mailing program can consist of:

- Just launched the application
- Just received an e-mail
- Replying to an e-mail
- Composing a new e-mail
- Just sent an e-mail
- Searching for an e-mail in the Inbox
- Adding a new mail account
- Adding a new messages filter
- Editing the program's configuration.

Each event should have a couple of questions assigned to it. The questions should relate to various UX and usability measures, such as effectiveness, efficiency, satisfaction, accessibility and safety [14]. A set of questions for an event in this study can look like following:

1. Does the software make your work easier?
   (Neg. Not at all; Pos: Much easier)
2. Are you satisfied?
   (Neg. Frustrated; Pos: Satisfied)
3. Is the use of the software enjoyable?
   (Neg. Annoying; Pos: Pure fun)

There can of course be many other studies, where the researchers can make good use of Crowdpinion. It all depends on the design of the study, types of questions, frequency of the events, length of the study, selection of the participants etc. As long as the event-contingent Experience Sampling Method is suitable for the study, the study should be possible to conduct effectively with Crowdpinion.

1.3. Challenges

There are always some challenges when designing and developing a software system, especially if the system needs to serve a research purpose. Out of the range of conceptual, technical and design issues that I had to face, I have chosen the two that had the largest impact on my work on Crowdpinion and described them below.

1.3.1. Responders' motivation in long-term studies

The studies in Crowdpinion are supposed to be conducted over a long period. This would bring many benefits to the researchers, because they would receive big sets of responses that can be analyzed in order to find trends in different moments and different events. However, in order to obtain loads of valuable data from users in long-term studies, we must have means to motivate the users to use Crowdpinion regularly and respond to questions whenever an event occurs. It is quite easy to imagine that after a few first days of moderately high interest in the study, the users' engagement will start to drop. It is a serious issue in a system like Crowdpinion, because if the users give less and less answers every day during a study, the distribution of responses in the whole study will be largely distorted and therefore less valuable.

Solution: Facing this issue, I have included some elements of gamification in the tool. Gamification is a technique of including the game elements in a non-gaming context for purposes other than just pure entertainment [11]. I have used a leaderboard of the most active participants and some features that can be unlocked by responding to questions. These elements are based on the participants' curiosity and the will to be better than others.

1.3.2. Making it easy for the responders

Apart from lack of motivation and engagement mentioned in the previous section, there is another problem involving participants in studies. The studies in Crowdpinion are supposed to obtain information (opinions, feelings, reactions etc.) about some events or actions that happen to users in during their day or in some precise periods. The difficult part here is that the users should use the app and respond to the questions shortly after the event. This usually means that the users would be using the app when they are busy.

Solution: In order to make responding to questions possible in such busy workday context, the workflow in Crowdpinion must be short, quick and simple. I have therefore reduced it to three simple steps:

1. Select a study
2. Select an event
3. Respond to questions

The user interface must be as well designed in a way that makes it most usable – the user should know immediately what they need to do in the app, because there is no time for confusion. The steps listed above are linear – the user goes forward without having to go back or thinking about alternatives. The requirement that I have set is that the users should be able to react to an event (respond to its questions) in less than 5 seconds.

1.4. Organization of the thesis

This master thesis contains five chapters including introduction. The content of the next four chapters is organized in the following way:

- Chapter 2: *Context and state of art*
  In this chapter I analyze the current state of research in the key concepts of the context of this thesis: obtaining opinion, crowdsourcing, gamification and UX evaluation.
- Chapter 3: *Crowdpinion*
  In this chapter I describe Crowdpinion as a software tool. I list the requirements, explain the architecture, implementation and the development process.
- Chapter 4: *Evaluation and lessons learned*
  In this chapter I present the results of the three studies that I did to evaluate Crowdpinion.
- Chapter 5: *Summary and future work*
  In this chapter I summarize the entire work on my thesis and discuss the possible further developments of Crowdpinion.

2. Context and State of the Art.

In my thesis work, I have researched several areas that together make the knowledge base for designing a surveying system for studies on momentary opinion.

The Experience Sampling Method has been the first thing that I investigated and that has become the methodology used in the tool. The set of various methods of asking questions together with advantages and drawbacks of each one has been a valuable input to the design of Crowdpinion.

I have also looked into the domain of Crowdsourcing, which is present in long-term studies. Crowdsourcing without monetary incentive is often done with support of gamification and it has been so in this case.

In the last part of this section, I tell about my theoretical research in momentary UX evaluation methods, which used to be the core interest in Crowdpinion and is still one of its primary areas of application.

2.1. Obtaining Opinion

It is rather easy to ask a person a single question. It gets much more complicated when we need to ask the person to express their feelings or opinion shortly after some specific event or in a busy situation. It is even more challenging when we want the person to respond regularly in a period of several weeks. In this section, I describe the Experience Sampling Method that gave the foundation for Crowdpinion. I also include the results of theoretical research in ways of asking questions that I did as a part of an essay.

2.1.1. Event-contingent ESM

The Experience Sampling Method (ESM) [15]is a research method designed by Larson and Csikszentmihalyi. It is primarily intended to capture the subjective experience of individual people in the context of everyday life. The authors emphasize the “ability to capture daily life as it is directly perceived from one moment to the next, affording an opportunity to examine fluctuations in the stream of consciousness and the links between the external context and the contents of the mind” [16] as the most important advantage of their method. It involves self-reporting, meaning that the participants of the ESM studies report their state of mind unsupervised in some moments – in regular intervals, at particular events or when a special signaling device triggers it. This is another advantage of the method, because it allows the researchers to obtain information about the participants’ activities that are not public (e.g. happen at participants’ homes). Furthermore, when the researchers are not physically present when the participants respond to questions, they do not bias or influence the responses with their presence.

Experience Sampling Method studies investigate the momentary feelings, opinions, experiences etc. This means that the questions should be answered during an event or immediately after it. Questions like “Are you happy now?” asked when a person is e.g. at work are suitable for ESM, while questions like “Are you generally happy at work?” or “Were you happy at work last week?” are not. The design of the studies and analysis of the results can include the person-centered and situation-centered questions. The first group focuses on how participants from different groups (classified by age, origin, background etc.) experience the same situations. The other one analyses the how the experience of individuals changes in various situations [16].

There are three general classes in terms of when the participants should respond to the questions:

- Interval-contingent sampling;
- Event-contingent sampling;
- Signal-contingent sampling.

The **interval-contingent protocol** involves experience sampling (responding to questions) in regular intervals of time or at the same fixed times every day. For example, researchers might want to find out how certain feelings (e.g. happiness, motivation, tiredness, will to interact with coworkers) change during days at work. They can then ask participants to respond to a set of questions every day from Monday to Friday at 8 am and 4 pm.

It is usually (unless the intervals are really short) the least burdensome study type, because the participants know how many times and when they need to report. It is however only suitable for studies that want to analyze participants reactions to events that are regular and quite long. It is easy to imagine that the regular responding times will not overlap with some short or irregular events.

In the **event-contingent ESM studies**, the participants are instructed to respond to the questions every time when one of events specified in the study occurs during their day. Studies that aim to investigate momentary opinion related to situations rather than times of the day (e.g. what do people think about when they use public transport) can implement this protocol.

This protocol is quite effective, because it guarantees that all (or more practically speaking: most) of the events of interest will be covered by participants’ responses. However, there is a risk that it will be too burdensome when the events occur very often. In a study where the initiative to respond the questions at particular moments comes from the participants, the questions will probably be left unanswered if responding to them is a burden.

The **signal-contingent protocol** is based on use of a signaling device. Originally, the researchers were using pagers, later palmtops and similar devices. Nowadays smartphones seem ideal for this purpose. In the signal-based studies, the study participants are asked to answer a set of questions (usually the same set included in an Experience Sampling Form) as shortly as possible after the signaling device gives a signal

– usually a sound or a vibration [16]. This protocol is suitable for researchers that want to ask questions often and in irregular moments, but without depending on events – for example to sample people's happiness in random moments, collecting ten samples every day.

As the previously described two approaches, also this one has its pros and cons. The advantage is undoubtedly that with the signals we can be quite sure that the participants will remember to respond to the questions in the moments when we want them to. On the other hand, the "beep" from the device can come in inconvenient moments, e.g. when the participant is in a long business meeting. In such situation, the beep itself is a problem and moreover the participant is unable to respond to the questions in the following hours.

The Experience Sampling Method has some pitfalls. It is quite burdensome for the participants, when they have to respond to questions many times a day over many days. As pointed by Scollon et al. [17], this leads to a risk that only some types of people agree to participate in ESM studies – for example, would lazy or busy people want to participate? This leads to group of responders potentially not reflecting the entire population.

Long-term studies can lead to problems with quality of data, which can decrease over time. After some days some participants can lose motivation and interest in the study and respond to less events or respond outside of the sampled time – e.g. respond to all of the day's questions in the evening, while it is important that they respond soon after the evaluated situations. Scollon et al. [17] say also that even if ESM allows reducing the memory biases, self-reporting is still biased by personal issues (social acceptance, cultural norms) of the participants.

The ESM has originally been created to detect *flow*. This is why at the beginning of my work on this thesis I was planning to create a tool focused primarily on the studies that would detect this psychological phenomenon. The name of "Flow" has been given by Mihály Csíkszentmihályi to the mental state of *optimal experience*. It is *"a subjective state that people report when they are completely involved in something to the point of forgetting time, fatigue, and everything else but the activity itself. It is what we feel when we read a well-crafted novel or play a good game of squash, or take part in a stimulating conversation. The defining feature of flow is intense experiential involvement in moment-to-moment activity. Attention is fully invested in the task at hand, and the person functions at his or her fullest capacity."* [18].

The state of flow is an important factor for both productivity and enjoyment. My Supervisor and I found it interesting enough to orientate the first moments of the thesis work around this concept. Then we noticed that in context of software users, the optimal User Experience can also lead to flow. This gave the direction to my further work.

2.1.2. How to ask and solicit answers to questions

At the beginning of my thesis work, I have written a Systematic Literature Review (SLR) [7] of momentary User Experience (UX) evaluation methods. It was motivated by the original idea to make Crowdpinion primarily adjusted for UX studies. Eventually the tool has become more universal in terms of studies it can be applied in, but some of the results of that SLR work are still relevant and interesting.

While working on the SLR I have analyzed 41 papers about various UX evaluation methods. Aside from the systematic review, I have extracted the information about what questions do the researchers actually ask, and what means do the responders use to reply. This resulted in an interesting comparison of several different methods.

Multiple choice

Questions for which the feedback is given by selecting one or more answers from the list provided by the study authors is among the most common methods. On one hand, it is good because giving feedback does not take too much time and the data is immediately organized into categories. On the other hand, predefined options do not give the participants possibility to express their experience in a more custom manner.

Open-ended questions

This kind of questions gives the participants a chance to write whatever they want, usually without limiting the length of the answer. Application of such technique has its downturns too, because giving feedback in such way is by all means distracting and takes a remarkable amount of time, so the participants are likely to lose both the flow state and the patience to give valuable answers. Nevertheless, this method can be effectively combined with multiple choice questions.

Sentence completion

Walsh et.al. [19] give an example of obtaining users' emotions by questionnaire consisting of sentence completion questions. The participants are asked to quickly and without much thinking complete a set of sentences, so that the answers describe, how the participants feel. The sentences can be like "The product feels…", "The main problem with the product is…". This kind of self-reporting makes it easier for the participants to express themselves, but at the same time they require the study conductors to classify the open answers to some categories to make the data more systematic.

Pick a word

Another rather quick method of obtaining feedback is used e.g. in the Microsoft Products Reaction Cards described in [20]. In this kind of method, the participants receive a large set of cards, where each card contains a name of an emotion (Calm, Friendly, Ordinary etc.) and are asked to pick the cards with words nearest to their current state. The original set prepared by Microsoft consisted of 118 cards. In this approach, it is important to have a good balance between the too low number of cards,

in which the users will not find the appropriate words and too big number that will make selection too time-consuming.

Emoticons

A study described by Meschtscherjakov et.al. [21] shows a another kind of non-verbal expression of feelings. The authors used a variation of the Experience Sampling Method where the participants basically choose one of 5 emoticons instead of choosing a textual answer or writing an open statement about their emotions. This approach reduces the effort required to answer the question to the minimum, which does not distract the user in their flow. The bad point is that the answers received in such study are extremely laconic and there is a risk that they do not deliver the full picture of the participants' state of mind.

Another non-verbal emoticon-based approach is mentioned as an alternative method provided by Microsoft in paper [20]. In this case the participants do not select an emoticon, but are given an emoticon or a photo of a face clearly expressing an emotion and are asked to state how much the presented feeling is similar to what they feel at the moment.

Photos, videos and phone's contextual information

Other ideas for reducing the participants' distraction are presented in [22]. In the presented approach, all the data about users' actions is collected automatically based on device's sensors (location, position, movement, time, open apps etc.) and the feedback reported by the participants is done by photos and videos captured by the device's camera.

2.2. Crowdsourcing

Along with the development of the Internet, an idea emerged, that instead of hiring expensive professionals or implementing complex programs to do some job, we can split the work into small and easy tasks and have it done by amateur internet users. This is the concept of *crowdsourcing*. I decided to include research in this field in my thesis, because in long-term experience sampling studies we do crowdsource the task of generating large amounts of data to the participants. Therefore, many characteristics and issues related to crowdsourcing are valid in Crowdpinion.

What is crowdsourcing?

In their research, Estellés-Arolas et al. [23] point out, that it is difficult to give an exact and universal definition of crowdsourcing. For example, some researchers accept the community-based services like Wikipedia or YouTube as crowdsourcing [24] while some other researchers do not (because the "product" – wiki articles are not used commercially) [25]. The aggregate definition by Estellés-Arolas et al. is worth citing as it covers most factors emphasized by various authors as properties of crowdsourcing:

"Crowdsourcing is a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcer will obtain and utilize to their advantage what the user has brought to the venture, whose form will depend on the type of activity undertaken." (Estellés-Arolas et al. [23])

Examples of crowdsourcing and studies about it

Let us have a look at some examples of crowdsourcing. Amazon Mechanical Turk [26] is one of the most popular platforms for crowdsourcing simple tasks (called Human Intelligence Tasks – HITs). In the Mechanical Turk, the requesters create HITs, which are simple tasks, often about classification (*"Classify Paragraph for a Legal Topic"*, *"Choose Image Subject Categories"*) or search *("Find the phone number of real estate market players"*, *"Find Contact Information for Article's Author").* The workers perform the tasks and if the requesters assess the result of the work as correct, the workers get paid, usually a fraction of a dollar. According to Amazon, there are over 500 000 workers from 190 countries and over 250 000 HITs at any time.

Crowdsourcing platforms like Amazon Mechanical Turk allow researchers to reach and recruit hundreds or more participants easily and at relatively low cost [27]. Because of this, there is indeed big interest in research on these platforms. Let us mention just a few examples of studies conducted using the Mechanical Turk. Heer and Bostock [27] conducted crowdsourced experiments on graphical perception and successfully replicated some earlier experiments. Kittur et al. [28] crowdsourced evaluation of Wikipedia articles and (after adjusting the tasks to prevent cheating) received results that were similar to evaluation by "experts" – Wikipedia administrators. Grady and Lease [29] crowdsourced tasks of assessing documents relevance. Shank [30] wrote about how the crowdsourcing platforms can be used in sociology and other social studies.

iStockphoto is mentioned by Howe [24] as one of the first remarkable applications of crowdsourcing. It is an online microstock photography provider, selling royalty-free photos. The clue is that the photos are not provided by professional photographers employed by iStockphoto, but by community of both professional and amateur photographers, who are paid a provision for every time when somebody purchases their photos in the website. Crowdsourcing of the collection of photos leads to iStockphotos being quite cheap for their customers. Daren C. Brabham conducted a study among 635 iStockphoto contributors asking them about their motivation. The results show four factors that were selected by more than half of the participants: financial incentive (89.8%), improving photography skills (79.1%), creative outlet (76.9%) and fun (71.9%) [31]. In subchapter 2.3 I talk more about motivation and how gamification fits into it.

Another example of crowdsourcing, that was analyzed by Brabham, is Threadless [32]. It is an online clothing company, which crowdsources the design work to a community of artists. The artists create T-shirt designs and submit them to the Threadless contest. If a design gets printed, its author receives a decent amount of money. The participants of Brabham's study indicated that the perspective of winning the money, obtaining recognition and feedback from other artists and possibility to improve their design skills are the key motivation factors [32]. These three seem to be important for the workers in creative crowdsourcing communities.

Both iStockphoto and Threadless have been quite successful businesses running on the crowdsourced products generation model. There is a study by Poetz and Schreier [33] suggesting that in the creative type of crowdsourcing, the quality of work done by the crowd matches or even exceeds the work by professionals. In their study, the authors gathered two sets of product ideas – one from professional designers and another from potential users of the products. Then they asked two experienced professionals to evaluate the ideas (without knowing whether they come from professionals or amateurs). Even though 18 out of 70 ideas submitted by users were filtered out as irrelevant or invalid, the remaining ones were in fact rated higher than the ideas from the professionals in terms of novelty, customer benefit and overall quality [33].

Issues related to crowdsourcing

The way in which crowdsourcing works can cast a major doubt – are the semi-anonymous workers in the crowd as reliable and productive as professionals? It is likely that the quality of crowdsourced work results will be worse [28]. It is not only because of lack of qualification of the workers. It is mostly because of presence of people who cheat – meaning that they fill the responses to tasks with any data just to earn money quickly. Eickhoff and de Vries [34] analyze several common ways of cheating in crowdsourcing and discuss the means that can be taken to detect it and reduce its impact on the results. The examples they give include selecting checkbox or radio options without thinking (e.g. always the first option, random option, all options etc.) and pasting the same string or fragments of an irrelevant texts in open text fields. These two obviously create lots of valueless data in the results of the crowdsourced tasks and at the same time are difficult to detect. The authors suggest several ways to prevent cheating: type of tasks (*malicious workers are less present in creative tasks*), control questions (questions that have one correct answer which can be given after reading the question properly) and filtering workers by origin or history (e.g. workers who *never respond to open questions* are likely to be the ones that do the tasks quickly and carelessly, thus with bad quality) [34].

The cost of crowdsourced work is definitely lower and because of that, the researchers usually have a financial reserve that can be spent on some sort of quality assurance – it can even be possible to ask multiple crowd workers to do the same task at little cost and compare the results to solicit the correct ones. Some crowdsourcing platforms, such as Amazon Mechanical Turk [26], have a system of reputation of the workers. Having this, we can have an opinion about reliability of specific workers. Kittur et al. [28] suggest also

that the right design of the tasks can largely improve the effort that the workers put in the task and therefore the value of the results.

Apart from the situations in which malicious users make the quality of results in valid studies a problem, there is another issue. The crowdsourced tasks themselves can be malicious when somebody makes use of the power of the crowd to achieve some malign goals. This is called *crowdturfing* [35]. Let us think about the mechanisms that have been created to protect websites from unwanted activity of automated robots. For example CAPTCHAs, that, because of their growing complexity and variety, are quite hard to pass automatically, are easy to solve by humans. Therefore, if somebody manages to embed CAPTCHAs into a crowdsourced task (e.g. “Copy the text from the picture”), they can break thousands of them very quickly using crowdsourcing. Apart from breaking CAPTCHAs, there can be many other crowdturfed tasks that in general are against the internet policies and etiquette, e.g. “Create a fake user account on Facebook and give me the credentials.”, “Write a bad comment about my competitor.”, “Find ten e-mail addresses for our marketing mailing.” etc. Even if the biggest crowdsourcing platforms disallow this kind of practices, there are growing dedicated crowdturfing systems [35] that support this unwanted branch of crowdsourcing.

Wisdom of the crowd

Wisdom of the crowd as a concept of obtaining some value from a big population of people is related to crowdsourcing and relevant to Crodwpinion. Wisdom of the crowd means that rather than asking a question to a single expert, we solicit the collective opinion of the crowd [36]. Wisdom of the crowd can be enabled by crowdsourcing [37] – the researchers can split a questionnaire into a set of crowdsourcable [34] tasks and obtain research data from the crowd. This (wrapped in the event-based Experience Sampling protocol [16]) is what Crowdpinion is doing.

2.3. Gamification

Workers in crowdsourcing (or in my case participants in long-term studies in Crowdpinion) need motivation to be active. There are obviously more possible means of motivation than just financial incentive described in the examples in section 2.2. Recently a concept of using fun as motivation in crowdsourcing in a similar way to how it is used in computer games is growing in popularity. This concept is called gamification.

What is gamification?

Gamification is a concept of providing entertainment instead of financial incentive as a reward for people who do some work in crowdsourcing [11]. It means that instead of paying the participants to do some crowdsourced tasks, we can add some game elements to a program so that the participants play a game while generating some data for us. The summary of the core elements of the gamification definition proposed by Deterding et al. include:

- *"the use* (rather than the extension) *of*
- *design* (rather than game-based technology or other game-related practices)
- *elements* (rather than full-fledged games)
- *characteristic for games* (rather than play or playfulness)
- *in non-game contexts* (regardless of specific usage intentions, contexts, or media of implementation). "[11]

Serious games

While in the classic gamification usually only some elements of games are used, there are multiple examples that use complete games to research purposes. They are generally called the serious games and can be defined as games "*played with a computer in accordance with specific rules, that use entertainment to further government or corporate training, education, health, public policy, and strategic communication objectives.*" [38]. An interesting example of a very effective serious game is the Galaxy Zoo [39], where the users are given small fragments of photos of galaxies and are asked to classify them. Only in the first year, more than 150 000 users contributed with an impressive number of 50 million classifications. Another successful example is Foldit [40] – a science discovery game which uses the players' ability to reason about 3D structures to determine shapes of proteins. In the first two years, there were 57 000 users working on a set of 600 protein structures. We can also mention two games being developed as parts of studies at Simula Research Laboratory: PictureSort and Picture Guess [41] [42]. In these games, the players get points for accurate and quick operation on photos (sorting and guessing of the content) and by playing the games they generate data for research.

Serious games are not always intended to generate data from human computation. Sometimes they are designed to raise social awareness and educate. Several examples of such games can be found in [43]. The paper describes three categories of serious games and gives interesting examples of games that won the contest at the *Games For Change Festival* in 2007. The first is the *awareness-raising games*, with an example of *Ayiti: The Cost of Life* [44], a game where players learn about the effects of poverty while managing the budget of a virtual rural Haitian family. The other category is the *transformation games*, which are oriented on important social issues and aim to transform the players' views on them. The award-winning game in this category is the *Peacemaker* [45], where the players are challenged to find a peaceful solution to the Israeli-Palestinian conflict. The third category is about the games that make a statement and inspire new thinking about important social issues. The example here is a game called *The Arcade Wire: Oil God* [46], where the player is the Oil God indeed and is given the destructive goal to double the oil prices using a combination of eight godly wraths. These games are high budget professional products supported by big organizations. There are however other examples from research background. An example of those is the Scientific Hangman [47], which is a simple mobile app, where the users are given points by responding to questions about pieces of research knowledge. Often giving the

right answer requires reading a paper's abstract and in this way the game spreads the scientific evidence among the general public.

Examples of gamification and related studies

Let us now have a look at some applications of gamification in its classical approach, where only some elements of games design are used, instead of entire games. The two most common gamification elements are **leaderboards** and **badges** [48]. A leaderboard is a list of users ranked by some criteria – usually descending by the number of points, which are granted to a user for performing the gamified activities. Therefore leaderboards usually show the most active users and allow each user to compare their progress with the results of the others [48]. Badges are "virtual goods" [49] that are given to (or unlocked by) users when they achieve something in the gamified system – e.g. reach some points threshold. Badges have several advantages: they *set goals* (e.g. "reach 100 points"), build *reputation* of a user and allow users to *identify with groups* of people who have the same badges [49].

Gamification in information retrieval is the domain that I have been closest to because of Crowdpinion and the GamifIR workshop where I have published my first paper [12]. There has been plenty of research in this field. The researchers have been analyzing methods of applying gamification to encourage users to generate structured information – e.g. to assess relevance of documents [50]. Researchers can also motivate people to do some actions for obtaining big amounts of information about the very usage – e.g. to learn how users browse search results [51]. Search itself can also be gamified [52].

To some extent we can say that the social networks use gamification to collect data – in networks like Facebook, Instagram, StackOverflow or YouTube people post and share texts, photos, videos etc. and are rewarded with "likes", "up votes", comments or any other kind of badges [49]. Often organizers of events announce a competition for the best photo from the event and using minimal cost and introducing the challenge to the event's participants, they obtain big number of photos [53].

Gamification has its applications outside of the software domain as well. Kamasheva et al. [54] describes how gamification is used by companies to motivate their employees to do some tasks or improve collaboration. In such case, gamification is of course a secondary motivational factor and it does not seem likely that it would ever replace financial incentives at work. There is also a very interesting application of gamification in car industry. Nissan Leaf is an electric vehicle. It is equipped with the Eco Mode software that tracks some parameters of eco-driving and uses them in a gamification context [55] – rewarding the drivers with immediate feedback and having a sort of social network with a leaderboard where the drivers can compare themselves to the others.

As pointed by Eickhoff [56], there are many more motivation factors other than just *money* and *entertainment*. Other elements of the motivation spectrum are for example *education* (people do some tasks because they allow them to learn something), *socializing* (working on something that enables interaction with other participants),

vanity (when doing something enables the participants to prove to the others and themselves how good they are) and *charity* (when the work has a purpose).

2.4. Crowdpinion as a UX evaluation tool

At the stage when the system (not called "Crowdpinion" yet back then) was planned as a UX evaluation tool, I have conducted a Systematic Literature Review [7] in momentary UX evaluation methods. It has been a very good way to obtain a big amount of well-structured information about the state of art of the domain. However, the SLR method is unfortunately very time consuming, so it has been impossible to do it for crowdsourcing and gamification.

Systematic Literature Review is an effective method of presentation of current state of research in given field. SLR is performed strictly according to a set of rules, including a predefined search strategy and inclusion criteria. [7] Making the literature review systematic helps in achieving more complete and more reliable results.

I have based my SLR on a complex search string, three research questions and a set of inclusion and exclusion criteria. The main steps of the work proceeding the writing of this paper were:

1. Specification of the research questions and the search string.
2. Search for papers in Google Scholar.
3. Papers selection in accordance to the criteria.
4. Data extraction
5. Data analysis

Research questions

The goal of a SLR is to gather a remarkable set of data that provides a quantitative answer to some research questions. In our SLR I wanted to learn more about the state of research in momentary UX evaluation and I have asked the following questions.

RQ1.1: What are the different ways of evaluating momentary user experience?

First, I wanted to know what methods or techniques can be used to evaluate the momentary User Experience. There are many methods, but not all of them can be applied to measure the short-term user's emotions. [9] For example, questionnaires that are broadly used in assessing the overall (long-term) user experience and are usually applied after the user has used the software, are not very likely to give good results when used to evaluate the momentary UX [57].

RQ1.2: Who are the participants of these evaluation methods?

It is quite clear that different groups of people react differently to software and technology in general. Therefore, each UX evaluation method can give different results

when used on different groups of users. I wanted to know who the users that participate in the studies described in the papers are.

RQ1.3: What types of systems or applications are evaluated?

According to Vermeeren et al. [9] more than two-thirds of UX evaluation methods are relatively independent on application type – they can be used to evaluate all or most kinds of software (desktop applications, websites, mobile apps etc.). This seems to be quite natural because even if the use of different kinds of software can cause different emotions, the very act of reporting the experience or observation is more or less the same. Nevertheless, I wanted to know what kind of software has been evaluated.

Literature selection, data extraction, data synthesis

During the entire process of literature selection, data extraction and data classification (synthesis), I have performed all our actions in a systematic way. Our motivation was to make the whole work repeatable, because I believe that this makes the results credible.

After many trial searches in Google Scholar and analysis of possible synonyms I have created the following search string:

("user experience evaluation" OR "user experience testing")
AND
"software"
AND
("momentary" OR "spontaneous" OR "episodic" OR "instantaneous" OR "short-term")
AND
human
AND
("emotion" OR "feeling" OR "reaction")

This means that I wanted to receive search results containing all papers about UX evaluation or testing in the domain of software, which mention momentary emotions, feelings or reactions of human users. In our opinion the search string allows us to assume that at least the major part of relevant research papers have been included in our review.

Having a good search string is not sufficient for getting a list of relevant research papers. Google Scholar as an automatic search tool will always return more results than needed. Therefore, I needed well defined inclusion and exclusion criteria to filter the Google Scholar search results and include only the relevant papers in our review.

I have decided to include only peer reviewed papers published in last 15 years, which describe user experience evaluation of working software. Working software status is an important criterion, because it is very unlikely to get any momentary UX in evaluation of system descriptions, prototypes or paper mockups. I have also decided to exclude all theses, books and technical reports.

Search results and analyzed papers

The search in Google Scholar, using our search string, returned 170 results. After applying the inclusion and exclusion criteria, many papers have been excluded for various reasons: some appeared to be irrelevant, many were MSc or PhD theses, which I excluded as not peer reviewed papers. After this step there were 66 papers left. During the next part of our work, the data extraction, I have excluded even more papers, which, apart from their partial relevance, did not contain valuable data in the areas that I was investigating. In the end there were 41 papers in our data extraction results.

Although the results presented a good variety of UX evaluation methods, only a little part of the described methods could be applied to evaluate the momentary User Experience. I have kept all the results in questions 2 and 3, while I have limited the results to 14 relevant answers for question 1.

RQ1.1: What are the different ways of evaluating momentary user experience?

In total there were 14 papers mentioning momentary UX evaluation methods. I have put the methods in five groups, as shown in Table 1. One of the papers ([58]) mentioned both the Experience Sampling and Think-aloud methods. It may be helpful to explain the three most common methods: Experience Sampling, Think-aloud and Product Reaction Cards.

The Experience Sampling Method (ESM) [15] is a research procedure which is commonly applied in a wide range of research areas – from psychiatry [5] to software engineering [59]. The main idea of the method in product evaluation is to ask participants to report their thoughts or feelings in random moments during the usage of the evaluated product. The characteristics of the method match exactly the requirements of momentary UX evaluation. It is not surprising then that this method is the most frequently used method in this field.

The Think-aloud method [60] is quite literally following its name. The participants are asked to speak out their thoughts and feelings while using the evaluated product (e.g. “Now I click this button because I want to confirm the decision.” or “Oh, this is confusing.”). This method requires a trained listener who is able to extract and record valuable data from the participant’s speech.

Table 1. Momentary UX evaluation methods

Type of method	Papers	Occurances	%
Experience Sampling	[22], [58], [21], [61], [62]	5	33%
Think-aloud	[63], [64], [58], [65]	4	27%
Product Reaction Cards	[66], [67], [20]	3	20%
Questionnaires	[68], [2]	2	13%
Biological reactions	[69]	1	7%

UX evaluation methods based on cards come in different variants (e.g. Emocards [66] or Product Reaction Cards by Microsoft [20]), but the main idea remains the same. The participants use a set of cards rather than speech to express their experience during usage of the evaluated product. Using cards results in more organized and probably faster feedback than in methods like the Think-aloud.

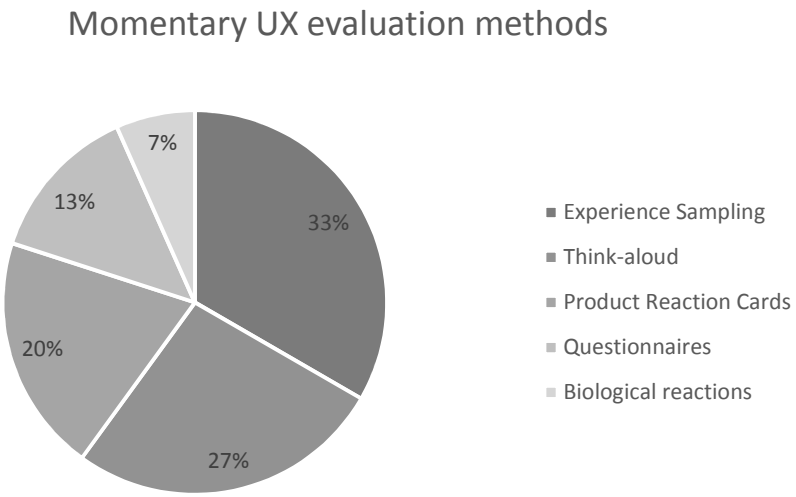

Figure 1. Momentary UX evaluation methods chart

RQ1.2: Who are the participants of these evaluation methods?

This question did not bring a big variety of answers. In all cases the participants were chosen from groups of potential users of the evaluated software. Some division can be made based on the level of experience of the users. In some case studies the participants were experienced, daily users of the software, in other studies the context was completely new to the participants.

In most cases the participants had at least some experience with the evaluated system and/or its context – e.g. participants who use smartphones every day were asked to test a new kind of app. Only few papers describe studies with completely inexperienced participants: [70], [66], [71], [72], [73], [74], [69], [75].

Obviously the authors of the studies had a purpose in selecting one group of participants or another. When the participants are familiar with the context, their reactions are based on the experience of the product itself. When the context is new, it is not so clear whether it is the product or the environment that triggers most of the feelings.

RQ1.3: What types of systems or applications are evaluated?

Originally the question was meant to be “What types of applications can be evaluated?”. However, during the data extraction I have not encountered any UX evaluation method, which would be limited only to one or two kinds of software. Apparently the same methods can be used to measure or evaluate users’ experience and emotions while the person is using a website, a mobile app or a mobile phone itself.

In order to at least show the distribution of presence of different types of systems in the research papers about UX evaluation, I have changed our research question to “What types of systems or applications are actually evaluated in the described studies?”. This gave us more diversified results.

The most commonly evaluated software types were: Web applications, Games (and other entertainment applications) and Mobile applications.

Table 2. Types of evaluated systems

Type of method	Papers	Occurrences	%
Web	[2], [63], [76], [77], [68], [78], [57], [79], [20]	9	24%
Games and entertainment	[80], [81], [77], [67], [64, 72], [65], [74]	8	22%
Mobile	[82], [22], [77], [67], [62], [83]	6	16%
Hardware devices	[19], [84], [85], [1], [86]	5	14%
Public services	[58], [87], [73]	3	8%
Augmented reality and AI	[88], [66], [71]	3	8%
Other	[70], [69], [75]	3	8%

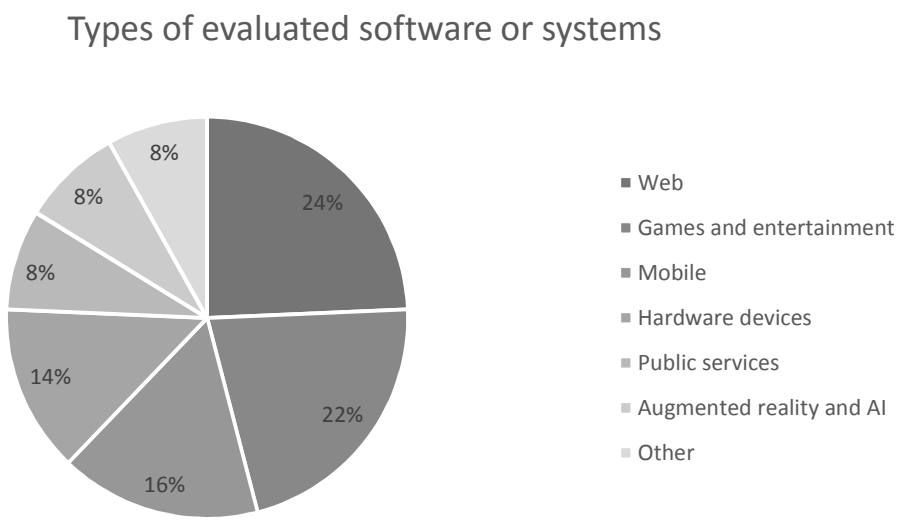

Figure 2. Types of evaluated systems chart

Discussion of the review results

In my systematic review, I analyzed the current state of research in the field of momentary UX evaluation methods. It differs from existing literature reviews and

expand the knowledge base by focusing only on momentary UX evaluation, instead of UX evaluation as a whole.

What I have found immediately after the data extraction was that the number of conference papers in this field is very low. Having 170 search results in Google Scholar and 66 included papers, I was hoping for a much more than the 14 papers talking about momentary UX. I have aimed my literature search directly into the momentary UX area and yet the number of papers was low. This makes me believe that even if I might have missed some papers in our work, the overall number of publications in the topic is low.

I claim that momentary UX deserves more attention. Even if in the end it is the overall user experience that shapes the opinions about the experience as stated by Ariely and Carmon in [89], it is vital to know how the experience changes over time during the use of a product. In the same paper, the authors claim that if the momentary experience changes more or less continuously from negative to positive over time, the overall experience will be positive. If it does the other way, the final effect will be negative. This is why in order to create products that are successful due to good overall UX, we do need to have measures to evaluate momentary experiences.

I have however observed that the earliest of the 14 analyzed papers that describe the momentary UX evaluation has been published in 2007 and the median year is 2010. This indicates somehow that the field is a relatively new thing that is likely to grow in the next years.

The relatively little volume of the data that I have managed to extract from the papers makes it hardly possible to make any interesting data-based conclusion. Out of the five methods of momentary UX evaluation that I have found in the literature, four were applied in evaluation of the whole range of software systems with different types of studies participants. The fifth method, the one that has been using the body's biological reactions (blink rate, the ratio between low- and high-frequency powers of heart rate (LF/HF), and maximum breathing frequency) [69] has been successfully used in a study, where the participants were a group of disabled people, barely able to communicate. This is an interesting example of an unusual method applied to unusual circumstances. The other methods (ESM, Think-aloud, Product Reaction Cards and Questionnaires) can be applied to more or less all software contexts.

Questionnaires were by far the most frequent of all (not exclusively momentary) UX evaluation methods that I have encountered in this study. The effectiveness of their application in the momentary UX domain is quite questionable though. Even if a questionnaire is short and is filled in immediately after the experience, by the time it is completed, the momentary emotion fades and changes [57]. Also using a using a fixed scale in suggested answers for the questionnaire's questions will sometimes subject the evaluation participants to framing effect.

3. Crowdpinion

In this chapter, I focus on Crowdpinion as a software system. I first describe the architecture of Crowdpinion to give you the overview of the system. Then I write about the requirements that gave the shape to Crowdpinion. In the next subchapters, I write more about the components being part of the system and in the last subchapter I describe the process of how Crowdpinion was designed, implemented and tested.

3.1. Architecture of Crowdpinion

Crowdpinion consists of several components, which together make it the complete surveying system. The components and the relations between them can be seen in Figure 3. There are two frontend components: the web panel and the mobile app. There are also three core elements on the server-side: the backend of the web panel (described further together with the frontend), the database and the web services used by the app. In this section, I describe the components briefly and in the next sections you can find the detailed descriptions of each of them.

Crowdpinion web panel

The web panel is the tool for researchers, which enables them to set up and manage studies. It is also used by the administrators to control the entire system. It is a CakePHP [90] web panel, where communication between users and database is implemented using the Model-View-Controller pattern [91].

Crowdpinion mobile app

The app is used by participants of studies. It allows them to quickly react to study's events and enjoy the gamification elements. It is a lightweight app created with the Ionic Framework [92] which includes AngularJS [93]. The app obtains studies' data and submits responses to database via the web services.

Database

The database used in Crowdpinion is a MySQL database [94] and consists of ten tables. It stores information about studies (events, questions, subscriptions, answers) and users (administrators, researchers and app users). The database is accessed by the CakePHP web panel and the web services.

Web services

The web services provide means for the Crowdpinion mobile app to access the database. There is a set of RESTful services written in PHP that support the mobile app whenever it needs to get a list of studies, events, questions or responses from the database or post the latest responses in the database.

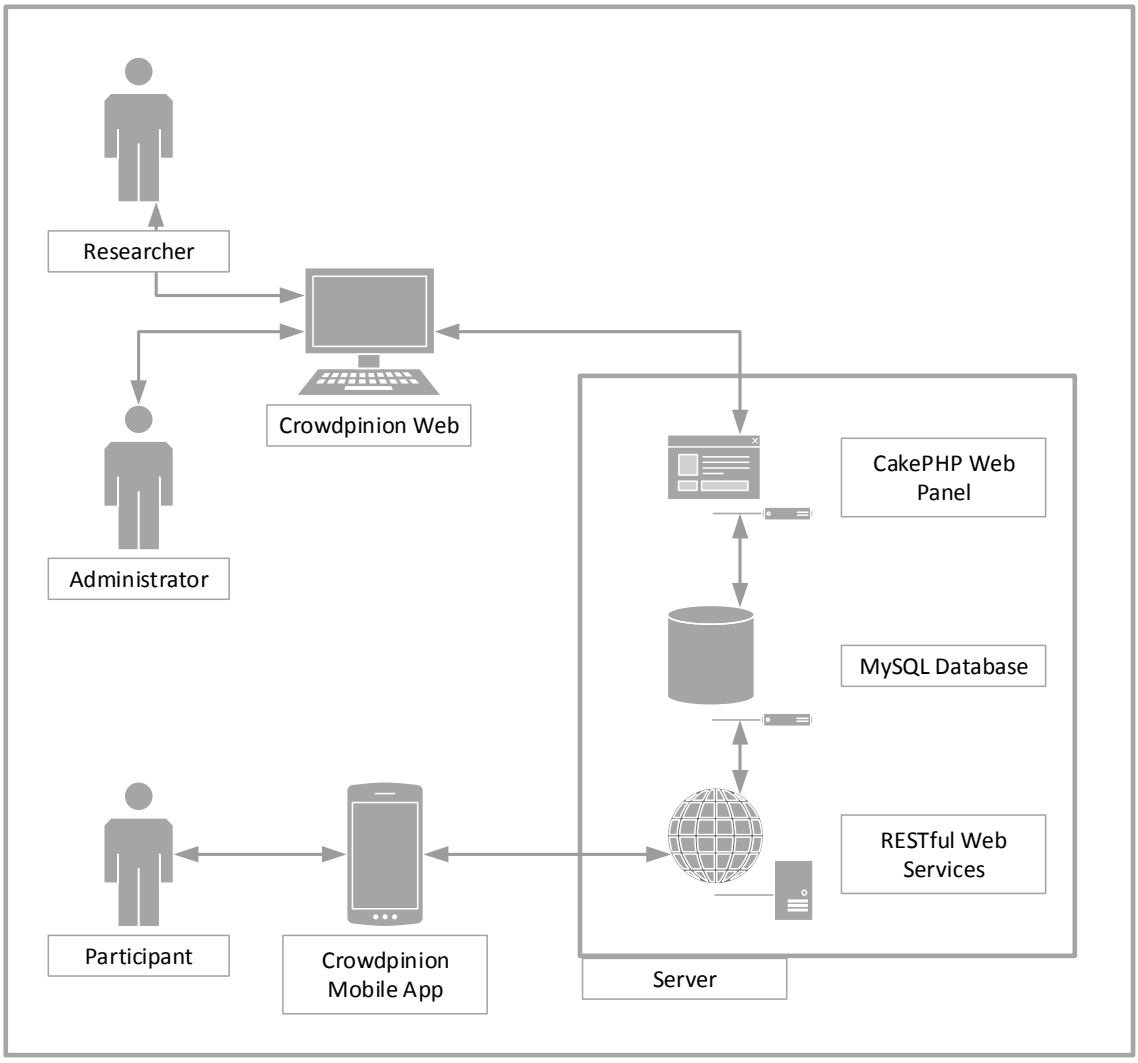

Figure 3. Architecture of Crowdpinion

3.2. Requirements

The requirements in this project have been being specified during several project meetings with my supervisor. In this very small team the Supervisor played the role of the Customer, who, based on his experience as a researcher, was talking about the desired functionalities of the software while I was the Business Analyst specifying the ideas into requirements and eventually the Developer coding the software. In the next sections, you can find the functional requirements specified in a minimalistic lean form of User Stories [95]. User Stories are a method of expressing requirements in a very concise way. They represent small pieces of requirements (business value) that can be implemented in a relatively short time – usually a couple of days. They are often expressed in a short template:

As a [role], I want [goal/desire] so that [benefit].

The role is one of the Actors in the system. The goal/desire is the actual thing that we want the system to do. The benefit is optional, but is often included for emphasizing the motivation behind a requirement.

3.2.1. Actors

There are three actors in Crowdpinion. Two of them - Researcher and Administrator – use the web panel. The Participant uses the app.

Participant

The Participant is a mobile user who wants to express his/her opinion about some topics by taking part in ESM studies. The Participant is contacted by the Researcher and invited to a study. A big population of active Participants (the Crowd) generates large amounts of valuable data for the studies.

Researcher

The Researcher is a person who conducts a study. They do not necessarily need to be professional researchers in an academic context. They can as well be for example an HR manager wanting to evaluate the attitude of the employees towards a new company policy. The Researcher sets up studies, invites participants, and analyzes the results.

Administrator

The Administrator is a person that has been granted full control over the system. They can do the same things as the Researcher but have additional administrative rights. The Administrator can moderate the work of Researchers.

3.2.2. User stories

The user stories are grouped by Actors.

Participant

All the user stories for Participant are related to the Crowdpinion mobile app.

1. As a Participant, I want to log in using Facebook so that the Authentication takes minimal effort.
2. As a Participant, I want to subscribe to studies so that I can contribute to research in the fields that affect me.
3. As a Participant, I want to see the list of studies that I subscribed to.
4. As a Participant, I want to see the details of a study.
5. As a Participant, I want to see the list of events for the selected study.

6. As a Participant, I want to respond to questions for the selected event so that I can express my momentary opinions.
7. As a Participant, I want to see the ranking of all participants so that I can challenge the others.
8. As a Participant, I want to see the summary of responses of all participants so that I can compare my opinions to the general trend.
9. As a Participant, I want to add my own questions to a study so that I can learn about what the others think about an issue that interests me.
10. As a Participant, I want to see the responses to the questions that I asked.
11. As a Participant, I want to unlock some features by answering questions so that I am motivated to give more responses.

Researcher

All the user stories for Researcher are related to the Crowdpinion web panel.

1. As a Researcher, I want to create an account in Crowdpinion.
2. As a Researcher, I want to log in with username and password.
3. As a Researcher, I want to edit my profile.
4. As a Researcher, I want to create a study.
5. As a Researcher, I want to edit a study.
6. As a Researcher, I want to delete a study.
7. As a Researcher, I want to create an event.
8. As a Researcher, I want to edit an event.
9. As a Researcher, I want to delete an event.
10. As a Researcher, I want to create a question.
11. As a Researcher, I want to edit a question.
12. As a Researcher, I want to delete a question.
13. As a Researcher, I want to invite Participants to a study.
14. As a Researcher, I want to see the responses so that I can analyze the results of the study.

Administrator

All the user stories for Administrator are related to the Crowdpinion web panel.

1. As an Administrator, I want to log in with username and password.
2. As an Administrator, I want to create a Researcher account.

3. As an Administrator, I want to delete a Researcher account.
4. As an Administrator, I want to moderate studies of other Researchers.
5. As an Administrator, I want to delete a Participant account.
6. As an Administrator, I want to have full control of the system.

3.2.3. Non-functional requirements

The non-functional requirements can also be expressed by User Stories. Several examples are listed below.

1. As a Participant, I want Crowdpinion to have high usability so that I can respond to a set of questions in less than 5 seconds.
2. As a Participant, I want the app to load its content within one second.
3. As a Participant, I want the system to make my responses anonymous.
4. As a Researcher, I want the system to keep my studies' data secure.
5. As a Researcher, I want the system to be available online at all times.
6. As a Researcher, I want the panel pages to load within one second.

3.3. Crowdpinion web panel

The web panel for the researchers allows them to set up and control studies. It is available online and the researchers can access it at any time.

3.3.1. Key functions

In this section, I describe the key functions and pages of the web panel. First I do it from the perspective of the researchers and then add the extended functionalities used by the administrators.

Log in page

All the web panel users use the same log in page. Then the system recognizes their role and loads the relevant panel. The new researchers can also create a new account. The new accounts are moderated by the administrators. The possibility to create new administrators' accounts is not open for everyone – only the existing administrators can add new administrators.

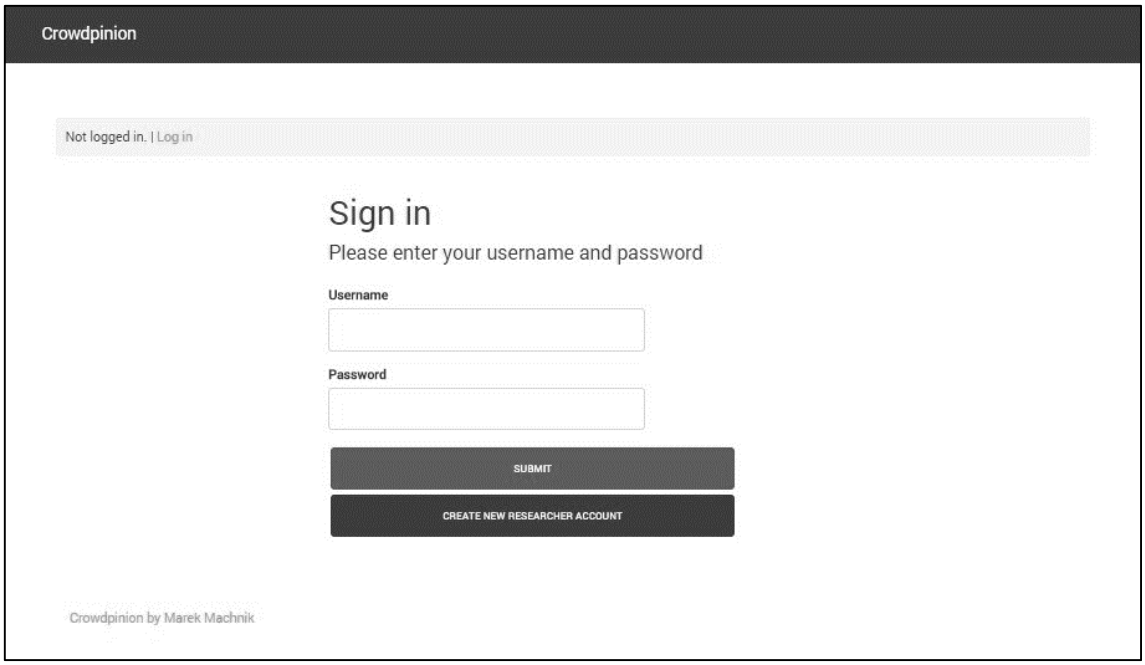

Figure 4. Log in page

After the username and password have been recognized as valid credentials of a researcher's account, the researcher is redirected to the home screen.

The home page (the studies page)

This page lists all the studies that have been created by the current researcher. A researcher cannot access the studies by other people. They cannot access the data about the users or system administration either.

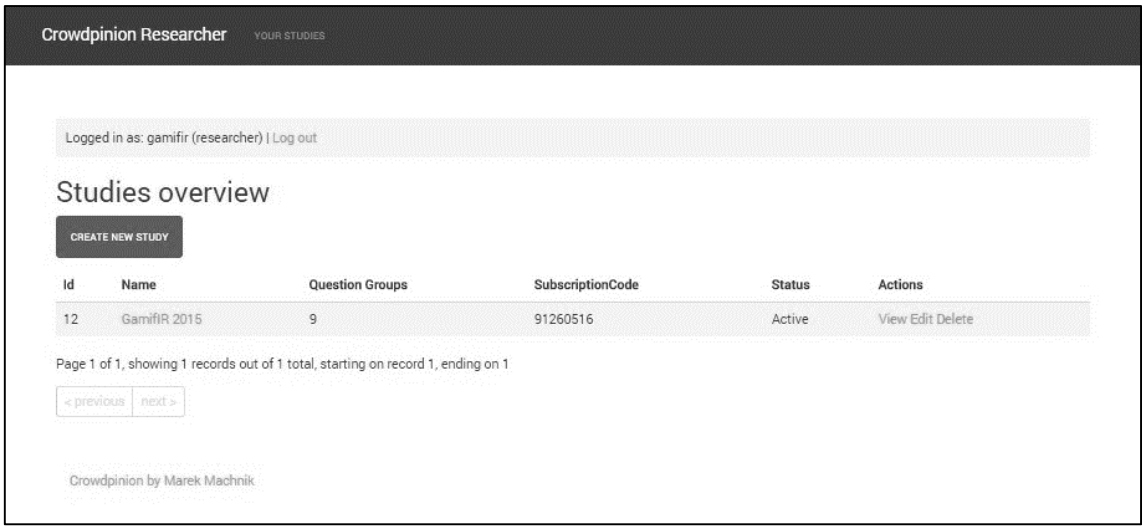

Figure 5. Researcher's home page

The researcher can view, edit or delete their studies. They can also add new studies.

Add a study page

In order to add a study, a researcher first has to specify its name and provide a short description. It is done in a simple form.

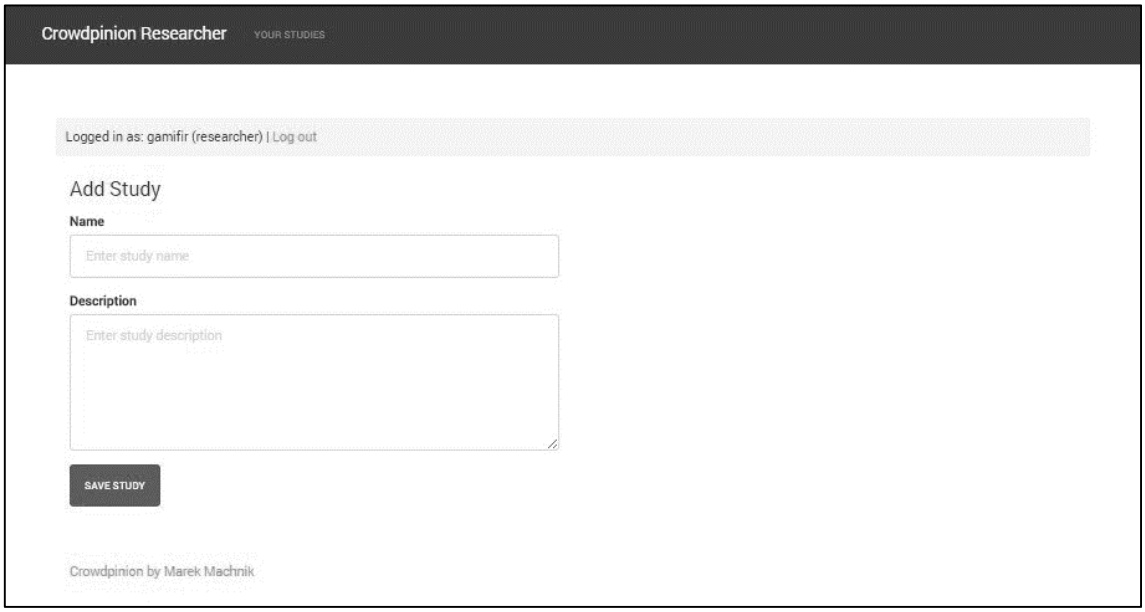

Figure 6. Add a study page

The study page

After creating a new study instance as described in the previous section, the researcher is redirected to the study view of the newly created study (Figure 7). This is the center of the study, where the researcher can set up the events and questions and continuously access the responses.

An important parameter of the study is the **subscription code**. It is a **randomly generated string of eight digits** that is unique for each study. While inviting the participants to the study, the only thing that the researcher needs to do, is to send the subscription code to the participants. Then the participants use the code in the app to subscribe to the study.

The events and questions can be seen and managed in the right column. Starting with the new study (Figure 7), the researcher first creates an event and then one or more questions to this event. An event has just a single field – a name. Questions are a bit more complex entities and they consist of the question text, labels for the positive and the negative answer and an optional comment. There is a dedicated form for creating and editing questions (Figure 8).

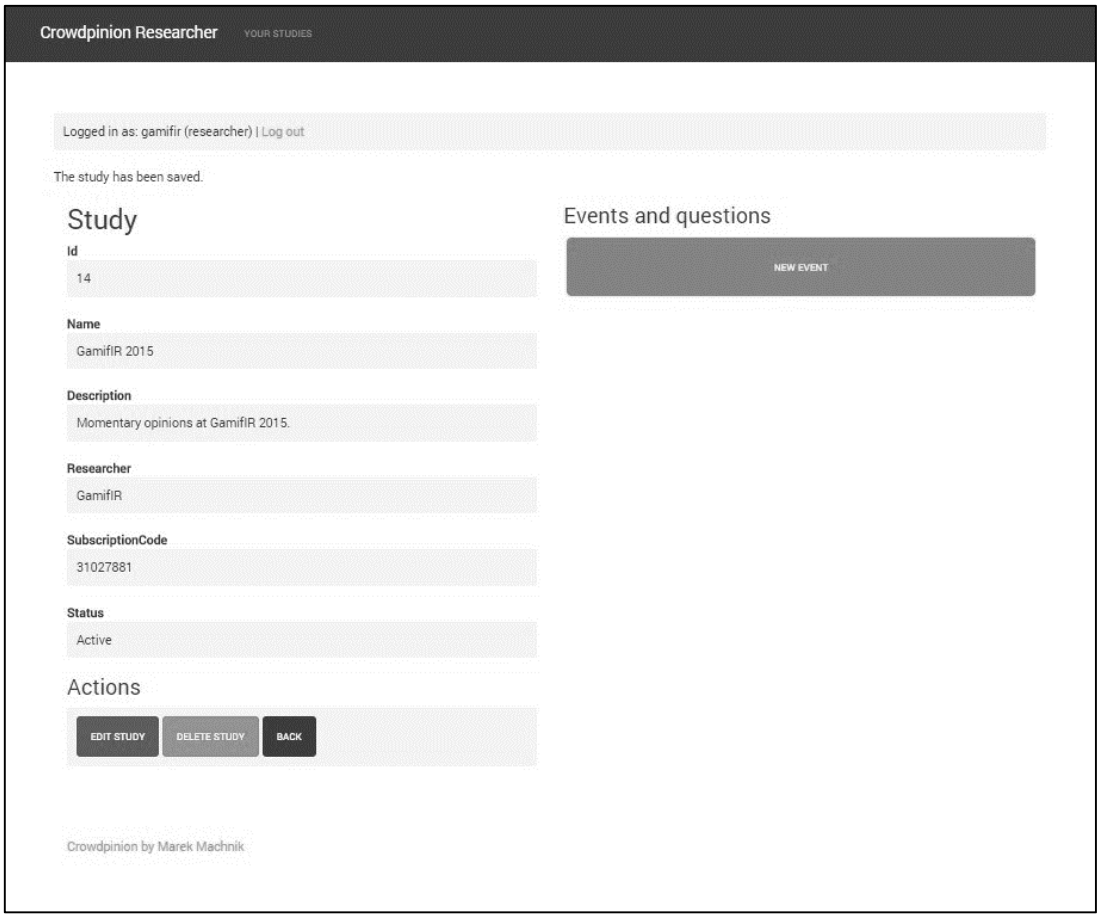

Figure 7. Study page - empty study

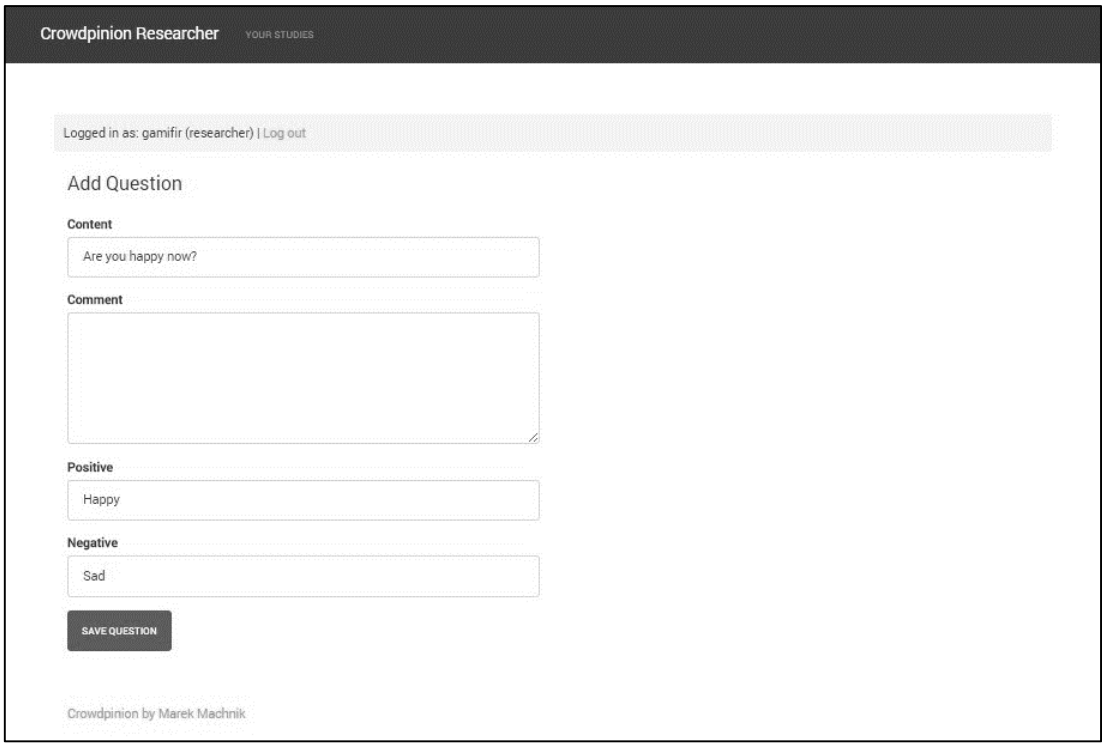

Figure 8. Add a question page

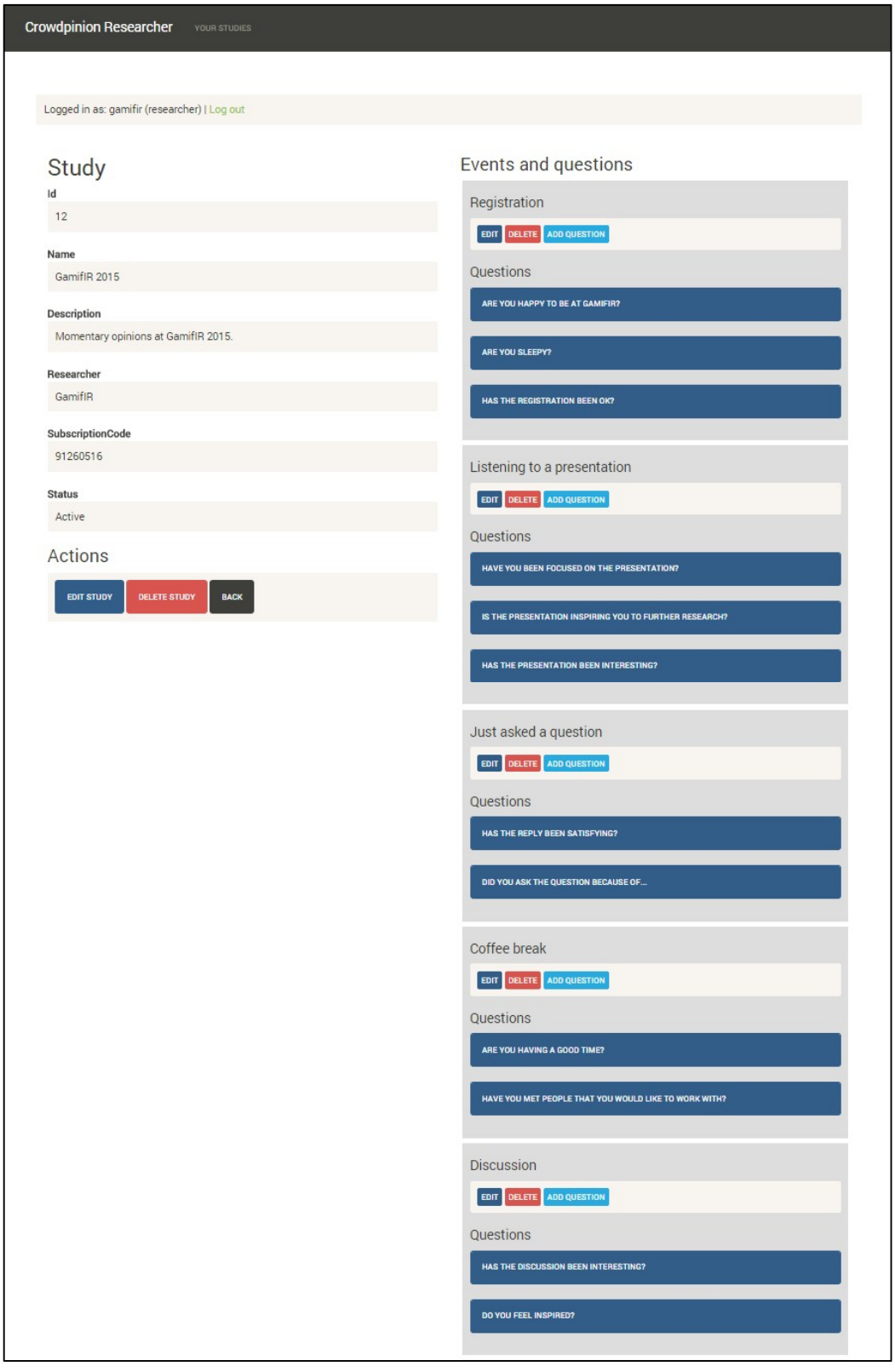

Figure 9. Study page - complete study

Once a study is filled with questions, it looks more or less as in the Figure 9. There is a long list of questions grouped in events. The researcher can edit the list of events and

questions whenever they need to during the study and the app will always load the latest version for the users.

The researcher can click the dark blue bar for selected question to expand the panels with responses (Figure 10). The way, in which the responses are presented in the current version of Crowdpinion, is usable but minimal. In the future versions I will extend the features for data analysis.

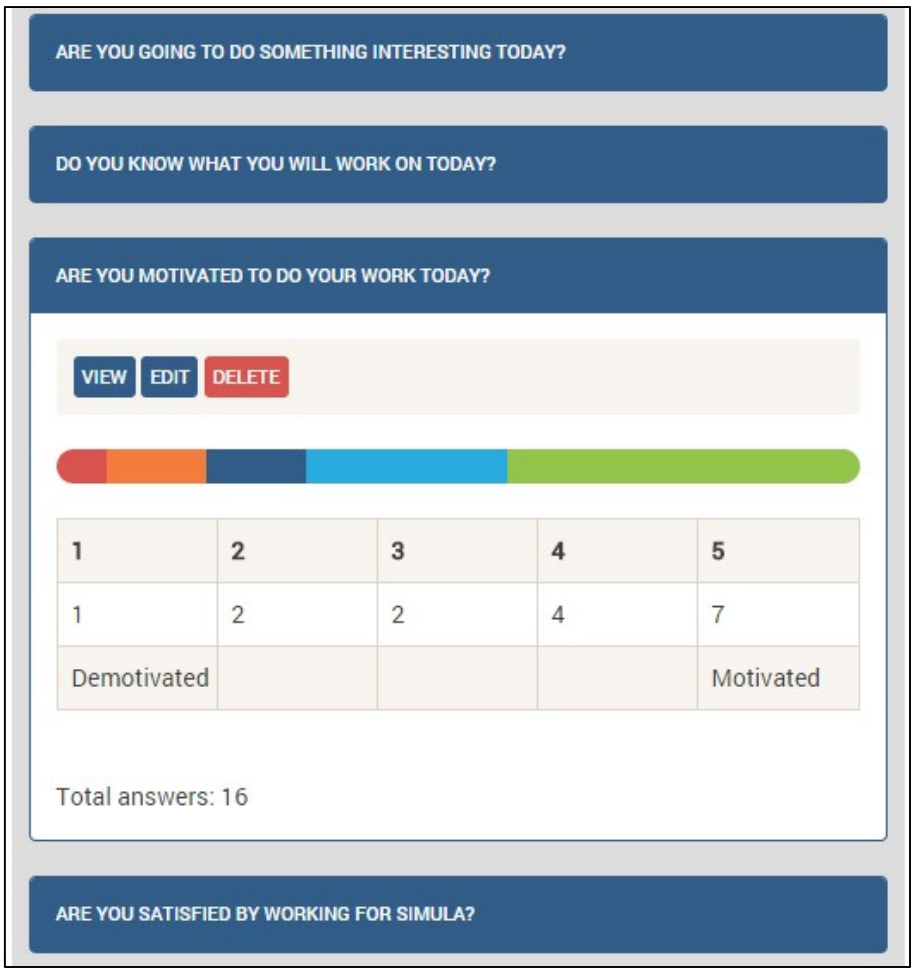

Figure 10. Study page - responses to a question

3.3.2. Extended administrative features

Administrators have access to all the functionalities that are available for researchers and to some additional features. They need to have full control over the system.

Manage studies

The administrators can see the list of all studies in the database. It is very similar to the list of studies that is accessible by the researchers, but is not limited to own studies. The

administrators can view, edit and delete any study. They can also create their own studies if they need to.

Manage the crowd

The administrators have access to a list of all users of the Crowdpinion app (the crowd members). They can see the users' data that has been obtained from Facebook as well as information about the studies that the participant has subscribed to and responses that they have given. Only the administrators can see this information. If it is still considered a confidentiality issue, I will limit the data that the administrators can access. Crowdpinion obtains only the very basic data from Facebook (name, e-mail address and user id) and no other personal information about the users is stored in Crowdpinion database.

Manage researchers

The administrators can see, edit and delete the accounts of the researchers. If a researcher needs assistance, they can change the researcher's password or some other details (Figure 11).

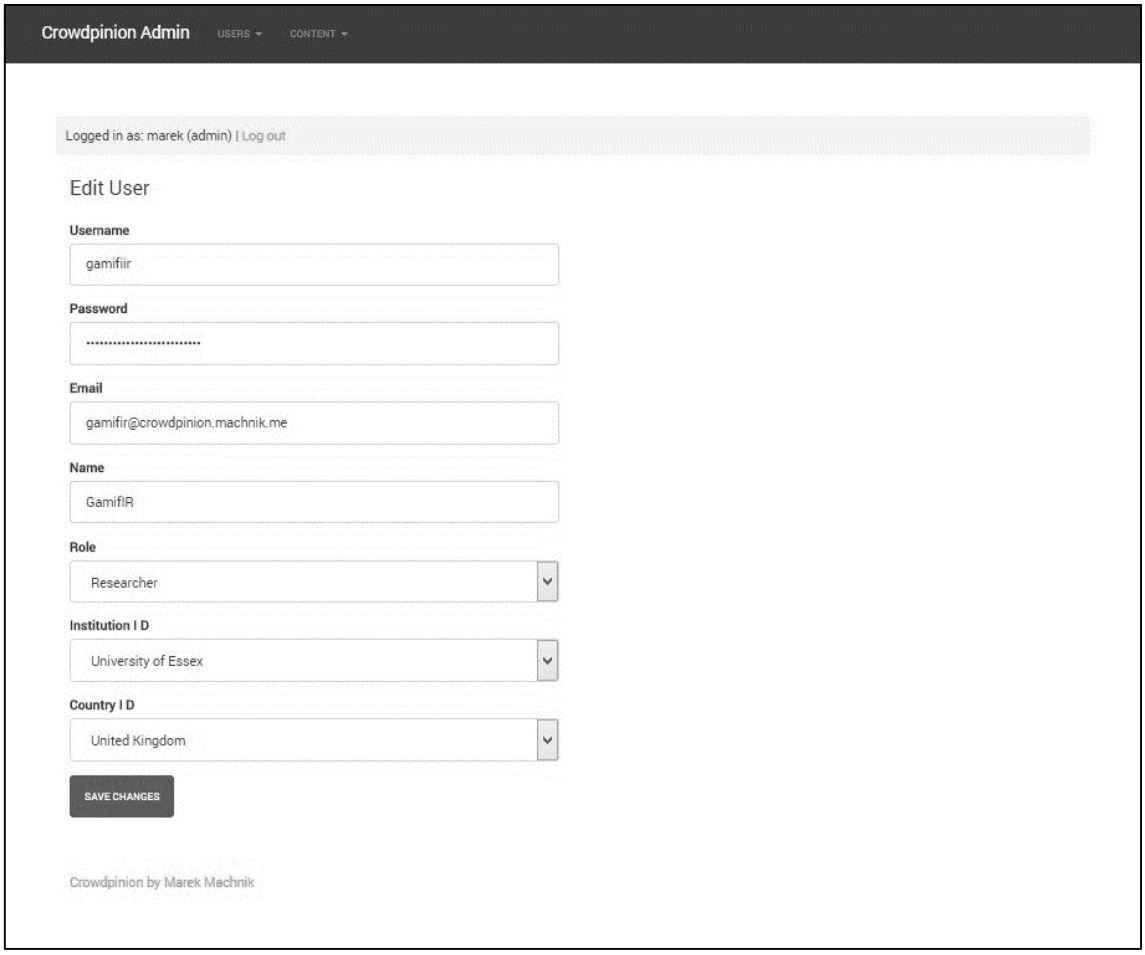

Figure 11. Manage a researcher profile

Manage institutions

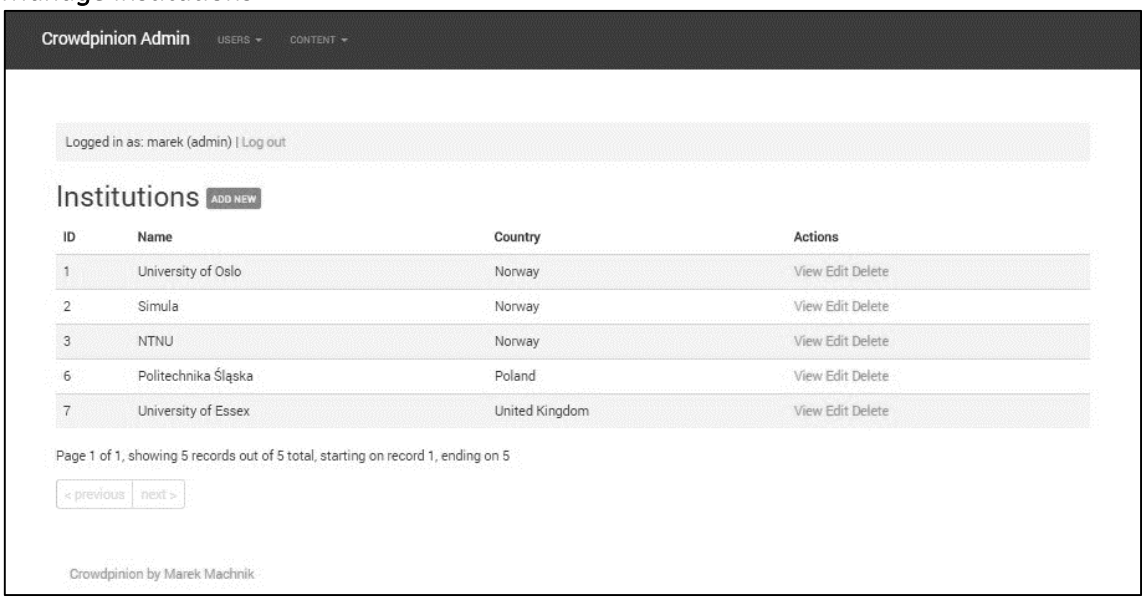

Figure 12. Institutions list

The institutions are just a property of the researchers – each researcher is assigned to an institution. The administrators can manage the list (Figure 12).

3.4. Crowdpinion mobile app

The participants of the studies in Crowdpinion use a mobile app. The app has been developed as a hybrid, which means that it can be built with Cordova [96] for different platforms without major modifications of the source code. Because of the limited access to Apple devices, the app has been built and used only on Android devices during my work on this thesis.

The primary goal of the app design has been to make it minimal. The app needs to be simple and quick to use. The participants can potentially use the app many times a day and because of it, I wanted to make it possible to react to an event in no more **five seconds**. The standard workflow in the app should be following:

1. Log in the app
2. Select a study
3. Select an event
4. Respond to questions

Even though the design could probably be more aesthetic, I think that I have managed to meet the requirement of simplicity and usability.

3.4.1. Key functions

In this section, I describe the core functions of Crowdpinion. As stated before, the app is minimal and simple, so there are only a few major screens.

Start screen

The start screen requires the user to log in (Figure 13). Even though the responses in Crowdpinion studies are anonymous, there needs to be an authentication mechanism to identify the user in the subscription mechanism and the gamification elements.

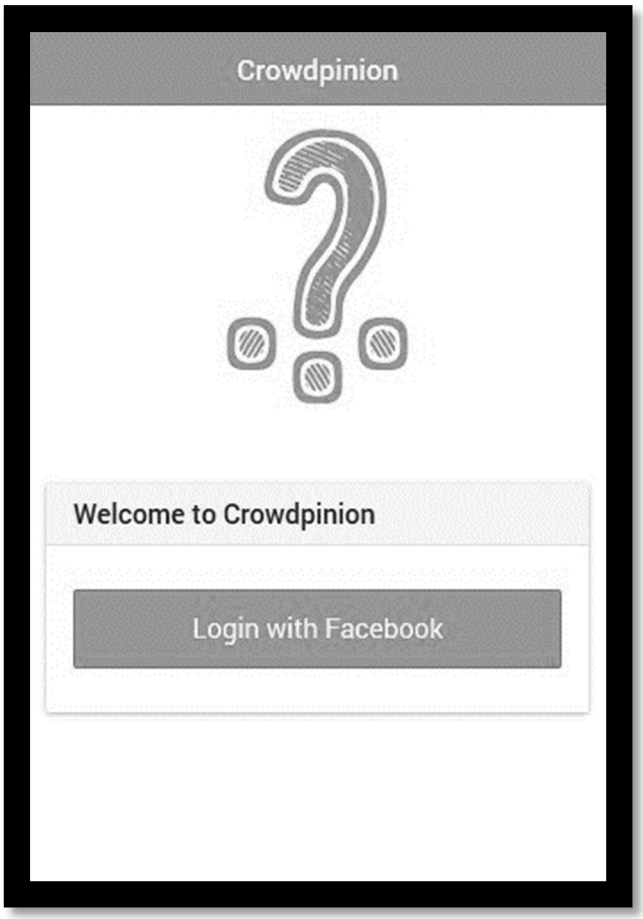

Figure 13. Start screen

I have chosen the Facebook login [97], which instead of requiring people to create new accounts in the app, allows them to use their Facebook credentials. It seemed to be the easiest solution for most potential users. However, this decision appeared to be surprisingly troublesome. Some users did not have Facebook accounts or did not want to use their data in an app, which they did not trust fully. One of the test studies' participants emphasized that she would not feel anonymous when logging in with her Facebook account. Openness to all users and the feeling of anonymity is vital for the app, so in the future versions of Crowdpinion, I will develop a different authentication mechanism – a custom authentication with nicknames rather than full user data.

Home screen

After logging in the system (or later after performing most of the actions), the user is redirected to the home screen. This screen (Figure 14) contains a list of all the studies that the user has subscribed to and a form that allows subscribing to new studies.

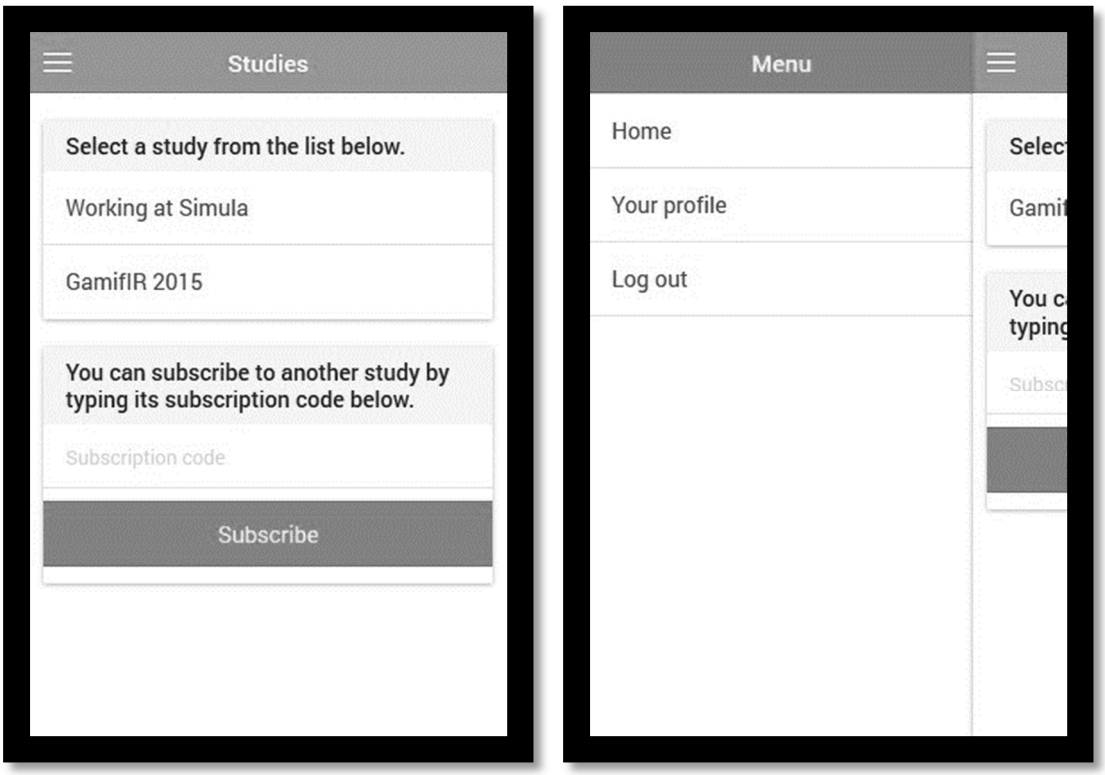

Figure 14. Home screen and main menu

The main menu is available at the left side by dragging the view right or by tapping the hamburger button. The content of the menu is minimal in the current version, but can contain more settings and features in the future.

Study screen

The study screen is another part of the standard flow (Figure 15). After selecting a study, the user sees a list of events available in the study. They also see three colorful buttons for the three gamification elements described in the Gamification section.

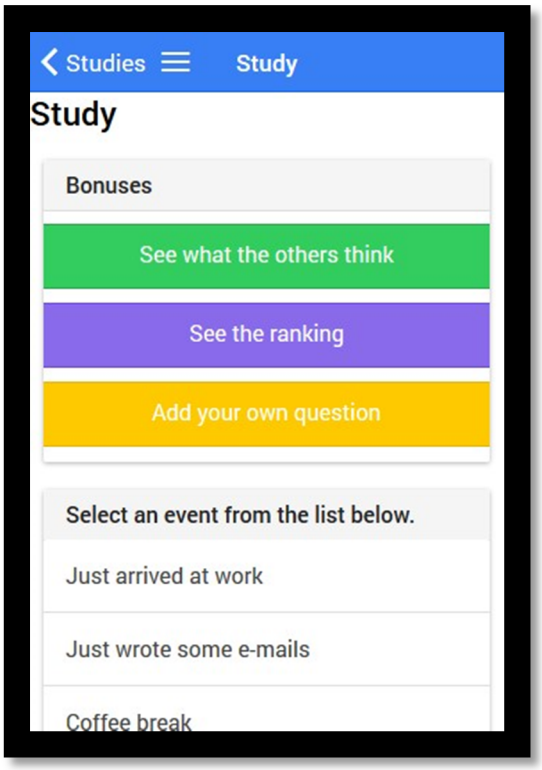

Figure 15. Study screen

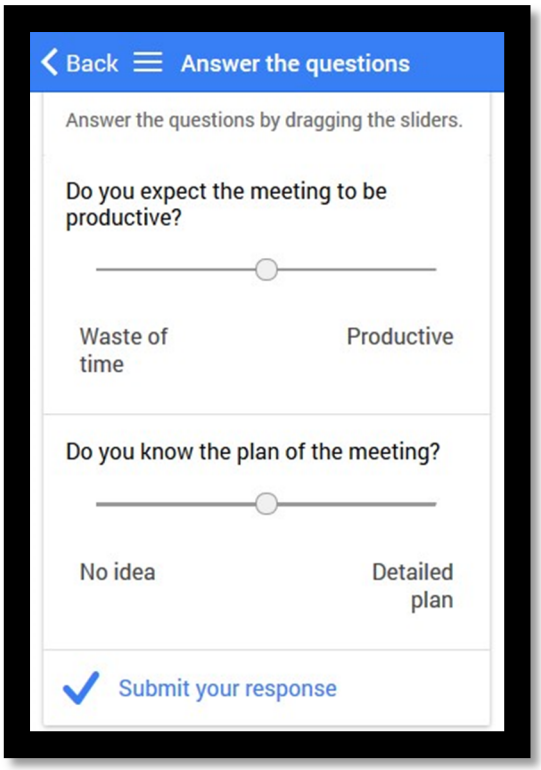

Figure 16. Questions screen

Questions screen

This screen contains the set of questions for the selected event (Figure 16). The question's content, positive and negative value are displayed together with a slider for each of the questions in the event. The user responds to the questions easily by dragging the slider right (towards the positive answer) or left (towards the negative). There are three intermediate possible answers between the negative and the positive, so there are five possible slider's positions in total.After the submission of the questions, the user is redirected to the home screen. According to the concept of the app, the user does not need to do anything else at this moment. However, if they want, they can still use the gamification elements. They should return to the app when another event happens.

3.5. Gamification elements in Crowdpinion

The users of Crowdpinion are presumably motivated participants. They subscribe to studies that investigate the issues that affect them or that they find interesting. Crowdpinion studies give them the opportunity to express their opinions and communicate what they feel about some issues. They want to contribute to assessment of the current situation and to possible improvements.

Still there is a remarkable issue with users' motivation and engagement present in this system. Experience Sampling studies can take many weeks and burdensomeness of repeatable responding to the same questions over a long period is often pointed as a disadvantage of the methodology [17]. In many cases, only the most motivated participants would stay active after the first week or two of a study.

In order to improve the participants' engagement in long-term studies, I have included some gamification elements in the Crowdpinion mobile app. They are based on building up the curiosity of the participants and introducing a bit of challenge between the participants.

3.5.1. Leaderboard

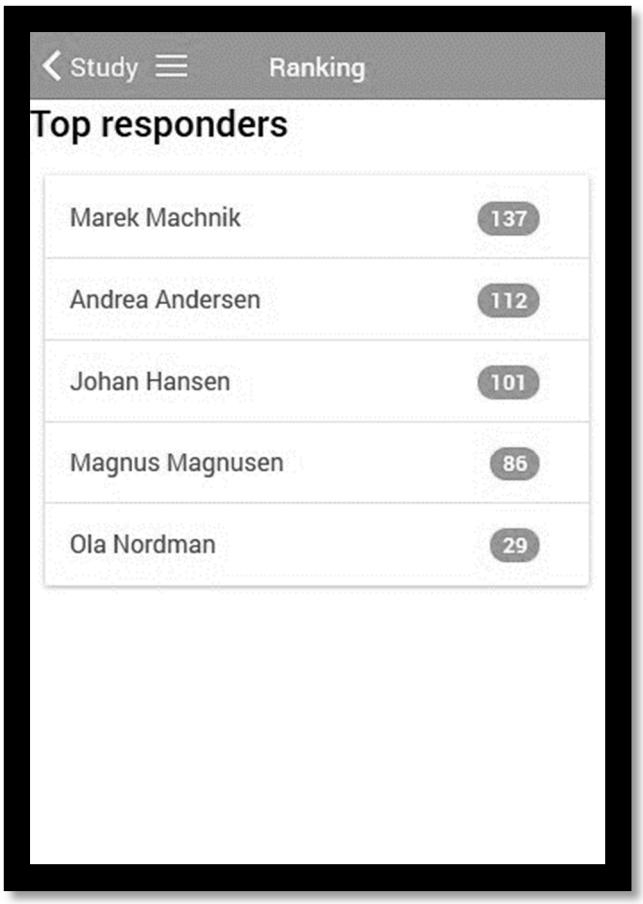

Figure 17. Leaderboard

The leaderboard is one of the most typical gamification elements. In Crowdpinion, there is a leaderboard for each study and it displays the participants, who have given most responses in the selected study (Figure 17). It is intended to motivate the participants to be more active in order to be better than the others.

3.5.2. Summaries of responses

One of the things that makes Crowdpinion special among the surveying systems is that the system shares the summaries of responses with the users (Figure 18). The participants can unlock the summaries of responses by active participation in the studies – this idea is based on the gamification concept of badges, but goes beyond it.

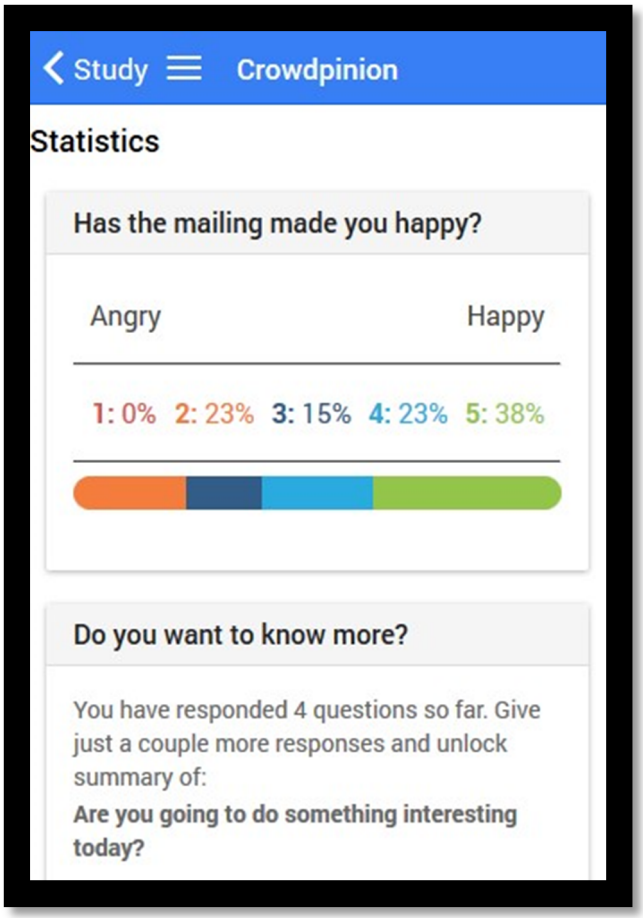

Figure 18. Summaries of responses

Initially (just after subscribing to a study) a participant can see the summary for the first question and an information on which question will be unlocked next. After responding to a set number of questions the participants unlocks the summary for question two and sees the information that the question three will be unlocked after some more responses. It is a gamification element heavily based on the participants' curiosity. I believe that the participants will want to know the results and how their feelings match the responses of the others.

3.5.3. Adding participants' own questions

It is another gamification element based on the concept of badges. For the participants that are interested in the topic of the study, this feature gives a lot of motivation to be active. They are given information that if they respond to a fairly big number of questions (e.g. 50), they will be able to add their own questions to the study (Figure 19).

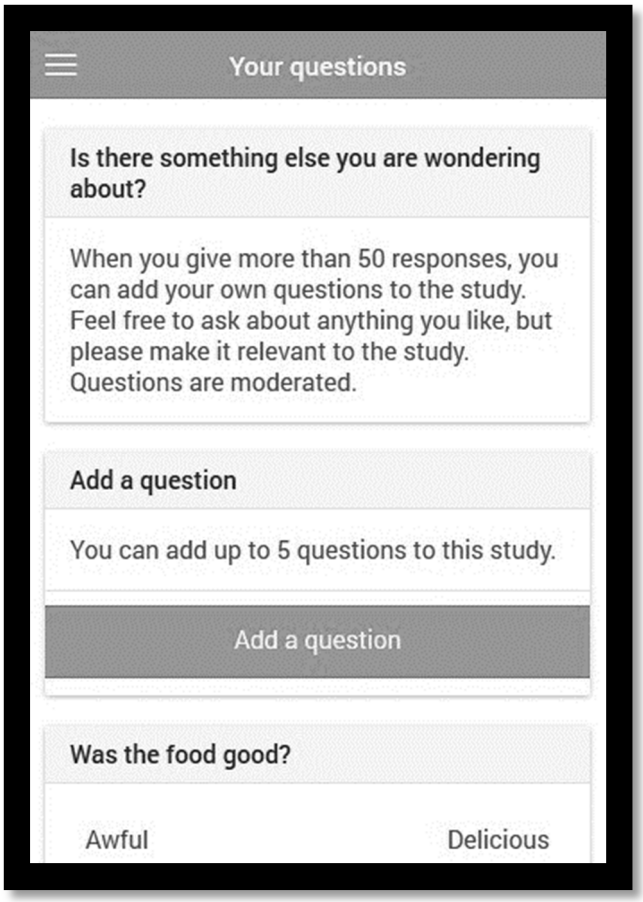

Figure 19. Participants' own questions

The participants can ask about anything they are interested in, but it is recommended that the question's topic is within the scope of the study. The researchers can moderate the questions added by the participants, so that the study is kept "clean". The questions asked by the participants are displayed in the app just as the original questions asked by the researchers. The participant that asked a question, can see the summary of the responses to it and the other participants can unlock the access to it afterwards.

3.6. Implementation details

In this subchapter, I give an insight in how I implemented the components of Crowdpinion using several popular technologies. I have included a couple of code samples that illustrate the characteristic fragments of the implementation.

3.6.1. CakePHP

I have used the CakePHP framework as a suitable solution to develop the Crowdpinion web panel. CakePHP is based on Model – View – Conroller, which is a design pattern introduced by University of Oslo professor Trygve Reenskaug [91]. It is a popular solution to making data accessible and manipulatable by users. It divides the functions of a software into three layers: model, view and controller [91] [90].

- Models represent the data. In CakePHP there is one model for each table in the database. A model describes the content of a table, constraints, relations between tables etc.
- Controllers handle requests from the users, connecting the users and the system. A controller prepares the data from a model to be displayed in views. If it receives a request from the user, it processes it, which often involves communicating with the corresponding model.
- Views are literally the elements responsible for displaying data to users. They are responsible for layout of data. It does not handle user's operation on data though – this is handled by the controller.

In CakePHP the models, views and controllers are linked together by a naming convention [90]. For example, in order to handle studies correctly and completely in CakePHP, the implementation should include at least:

- Database table: **studies**;
- Model class **Study** in **Model/Study.php** file;
- Controller class **StudiesController** in **Controller/StudiesController.php** file;
- View files **add.ctp**, **edit.ctp**, **index.ctp** and **view.ctp** in **View/Studies** directory.

Each view must have a corresponding method in the controller, but not every method in the controller needs a corresponding view – methods like delete perform an operation and then return to the view from which they were called. There can of course be more methods than just standard set of add, edit, delete, view and index.

```
public function index() {
 $this->Paginator->settings = array( 'conditions' => array(
'Study.ResearcherID'=> $this->Auth->user('id')), 'recursive' => 2);
 $this->set('studies', $this->Paginator->paginate());
}

public function add() {
 if ($this->request->is('post')) {
  $this->request->data['Study']['ResearcherID'] = $this->Auth-
>user('id');
  $this->request->data['Study']['SubscriptionCode'] = rand(10000000,
99999999);
  $this->Study->create();
  if ($this->Study->save($this->request->data)) {
   $this->Session->setFlash(__('The study has been saved.'));
   return $this->redirect(array('action' => 'view', $this->Study-
>field('id')));
  } else {
   $this->Session->setFlash(__('The study could not be saved.
Please, try again.'));
  }
 }
}
```

Figure 20. CakePHP controller methods

Figure 20 presents examples of two controller methods for studies in Crowdpinion web panel. The index() method the data to display in the table of researcher's own studies (Figure 5) and is in fact very simple. It obtains data about studies, filtered by the Id of the current researcher and copies the result array to the studies variable that can be accessed from the view.

```
<div class="col-lg-6">
 <?php echo $this->Form->create('Study'); ?>
  <fieldset>
   <legend><?php echo __('Add Study'); ?></legend>
   <h2>Settings</h2>
     <div class="form-group">
    <?php echo $this->Form->input('Name', array('class' => 'form-
control', 'placeholder' => 'Enter study name')); ?>
     </div>
     <div class="form-group">
    <?php echo $this->Form->input('Description', array('class' =>
'form-control', 'placeholder' => 'Enter study description')); ?>
     </div>
  </fieldset>
  <?php
   echo $this->Form->Submit(__('Save study'), array('class' => 'btn
btn-primary'));
   $this->Form->end( );?>
       </form>
</div>
```

Figure 21. CakePHP add study view

The add() method is also fairly simple. It obtains the data from the add study form in the view (shown in Figure 21 and in Figure 6), combines it with current researcher Id and

randomly generated subscription code and saves it in the database. If the save is successful, it also redirects the researcher to the view of the newly created study, so they can continue by adding events and questions.

3.6.2. AngularJS, Ionic, Cordova

The Crowdpinion app is a hybrid mobile app [98], which means that is has been developed using web technologies such as JavaScript, HTML and CSS and then built using Cordova [96], which wrapped the web application in native shells typical to mobile platforms. Because of this, even though the app is developed as a web application, the output of compilation is an .APK or .IPA file, which can be installed on Android and iOS platforms, distributed in the mobile app stores etc. Hybrid apps can easily access the features of the device, such as camera or GPS, which is important for the future development of Crowdpinion.

I have used AngularJS [93] as a core of the implementation of the app. Angular is an open-source web application framework, suitable for developing web applications using the Model-View-Controller pattern [91]. AngularJS, as opposed to jQuery [99], separates DOM (Document Object Model) manipulation from the application logic, while it keeps the access to the views easy by introducing the two-way data-binding. The two-way data binding is a mechanism that allows both displaying the data from the model in view and updating the model automatically when the data is modified in the view.

Let us have a look at an example of how I used AngularJS and Ionic in the Crowdpinion mobile app. Figure 22, Figure 23 and Figure 24 show the source code for the controller, the services and the view of the questions screen of the app ().

```
.controller('QuestionsCtrl', function($scope, $stateParams,
$localStorage, Services, $state) {
   Services.getQuestions($stateParams.eventId).then(function(data){
      $scope.questions = data.data;
      $scope.questionResponse= {};
   });

   $scope.submitQuestionResponse = function(questionResponse) {
      $questionResponse = $scope.questionResponse.resp;
      $user = $localStorage.getObject("user");
      for (var key in $questionResponse) {
         $formattedResponse = '{"Value":"'+ $questionResponse[key]
+'","QuestionID":"' + key + '","MemberID":"' + $user["id"]  %
100000000 +'"}';
         $response =
Services.insertQuestionResponse($formattedResponse);
      }
      alert("Response sent");
      $state.go('app.studies');
   };
})
```

Figure 22. AngularJS controller code for questions

The callback of the `Services.getQuestions(eventId)` method assigns the data (an array of questions) retrieved from the web services to a scope variable `$scope.questions`. The view can then display the questions using the Angular for-each loop `ng-repeat="question in questions"` and referring to each question field by `{{question.field}}`.

```
obj.getQuestions = function(eventID){
  return $http.get(baseURL + 'questions?id=' + eventID);
}

obj.insertQuestionResponse = function (questionResponse) {
  return $http.post(baseURL + 'insertQuestionResponse',
questionResponse).then(function (results) {
   return results;
  });
};
```

Figure 23. Fragment of AngularJS services code

When an user responds to the set of questions and confirms with the “Submit your response” button, the `submitQuestionResponse(questionResponse)` function is called. The function combines the responses with the data of the current user into a JSON [100] string and sends it to the web services [101].

```
<ion-view view-title="Answer the questions">
 <ion-content>
  <div class="list card">
   <div class="item item-text-wrap">
    <p>Answer the questions by dragging the sliders.</p>
   </div>
   <form ng-submit="submitQuestionResponse(questionResponse)">
    <div class="item question" ng-repeat="question in questions">
     <div class="item-text-wrap">
      <h2>BRACKETS question.Content}}</h2>
     </div>
     <div class="range range-positive">
       <input type="range"
ng-model="questionResponse.resp[question.id]" min="1" max="5" ng-
value="3">
     </div>
     <div class="row">
       <div class="col item-text-wrap">BRACKETS
question.Negative}}</div>
       <div class="col"> </div>
       <div class="col item-text-wrap" style="text-
align:right">BRACKETS question.Positive}}</div>
     </div>
    </div>
    <button type="submit" class="item item-icon-left positive">
     <i class="icon ion-checkmark-round"></i>Submit your response
    </button>
   </form>
  </div>
 </ion-content>
</ion-view>
```

Figure 24. AngularJS view code for questions

3.6.3. RESTful services

Crowdpinion mobile app uses the web services to communicate with the database. The services in Crowdpinion are a set of relatively simple RESTful web services. REST (Representational State Transfer) is an architecture for developing web services that are scalable, have high performance and are accessible via simple interfaces [102]. The web services in current version of Crowdpinion have been implemented following the services architecture presented in the tutorial [101] by Swadesh Behera.

As presented in the section about the app implementation, the client (the mobile app) communicates with the services in a simple way – for example `$http.get(baseURL + 'questions?id=' + eventID);` is enough to access the service that returns the set of questions for the given event. The web service [101] (shown in Figure 25) receives the request together with the eventId parameter and queries the database for all questions related to the event with given Id. If there are any results, the service puts them in an array, which is then converted to JSON [100] and sent back to the client that made the request.

```
private function questions(){
 if($this->get_request_method() != "GET"){
  $this->response('',406);
 }
 $id = (int)$this->_request['id'];
 $query="SELECT distinct q.id, q.Content, q.Negative, q.Positive
FROM questions q where q.GroupID=$id";
 $r = $this->mysqli->query($query) or die($this->mysqli-
>error.__LINE__);

 if($r->num_rows > 0){
  $result = array();
  while($row = $r->fetch_assoc()){
   $result[] = $row;
  }
  $this->response($this->json($result), 200);
 }
 $this->response('',204);
}
```

Figure 25. Example of a web service

It is similar in cases when the app needs to insert responses, user data etc. into database. Then the service receives data in JSON, parses it, inserts it in the database and reports the insertion status back to the app.

3.6.4. Database

Crowpinion utilizes a MySQL database. The web panel accesses it directly with the entire CakePHP application built on it (as described in the CakePHP section). The mobile app communicates with it via the web services. The database consists of 10 tables of different structures and purposes.

Figure 26 illustrates the structure of the database. More details about the structure of specific tables can be found in **Appendix E. Database details**.

Database diagram

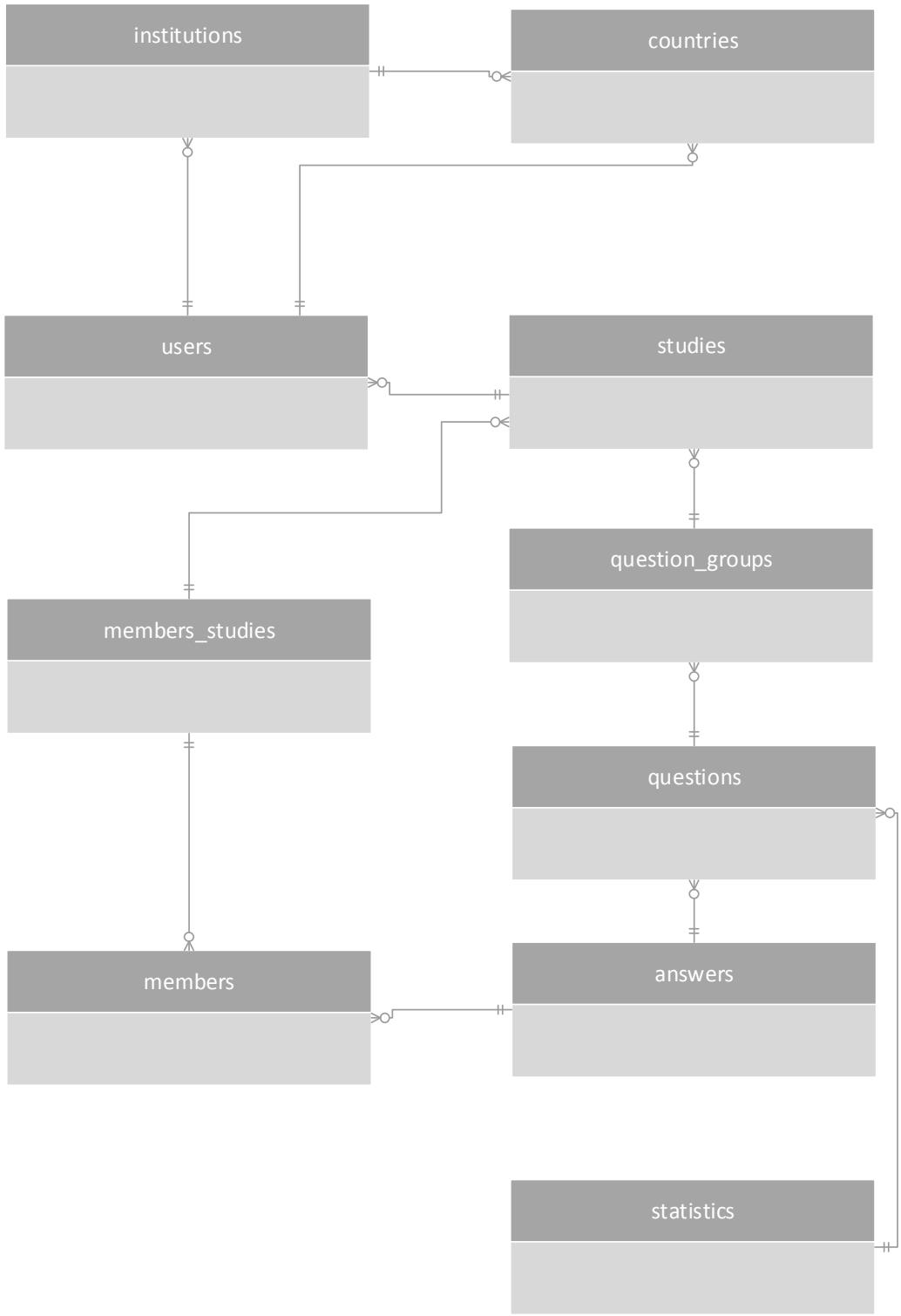

Figure 26. Database diagram

3.6.5. SVN Version Control

Since the beginning of the development work on Crowdpinion, I have been storing the source code of all components of the system in an SVN repository [103]. A repository is a storage location, which assigns a version to each file and contains detailed information on changes in the files. Every time when I managed to extend the existing code base with another working function, I tested the system for regression and I committed it to the repository. It is a very useful and important practice for several reasons.

- Possibility to revert undesired changes. The repository is supposed to store always a working version of the system. This means that if a developer makes some changes, which go wrong and the modified software stops working correctly, he can compare the modified files with the corresponding files in the repository and revert some or all changes.
- Storing the code in a secure location – SVN as a backup. Usually a repository is located on a remote server. This means that if the developer's computer breaks down, the code can be retrieved from the remote location.
- Collaboration. This feature I did not use in my project, because I was the only person working on the code. However, if multiple developers work on the same project, a repository allows them to synchronize the code and make sure that there are no conflicts.

While luckily I did not need to retrieve the entire code from SVN backup, I have used it many times to come back to the stable version after some faulty changes.

3.7. The process of creating Crowdpinion

In this section, I tell about how I have worked on Crowdpinion together with my Supervisor. I mention the requirements phase and the creation of mockups of the layout and flow. Further I write about the development and testing of the software.

The requirements for the Crowdpinion project have been shaped after quite a long time of conceptual work with my Supervisor. More information about the evolution of the ideas and the shaping of the requirements can be found in the previous chapters of this thesis.

When the requirements took shape of a more or less complete system, I have created a set of interactive mockups of the web panel. These mockups were an HTML website without a database. This means that it presented the layout and the flow, but all the data displayed in the screens was a sample data and it was impossible to add, edit, delete any items.

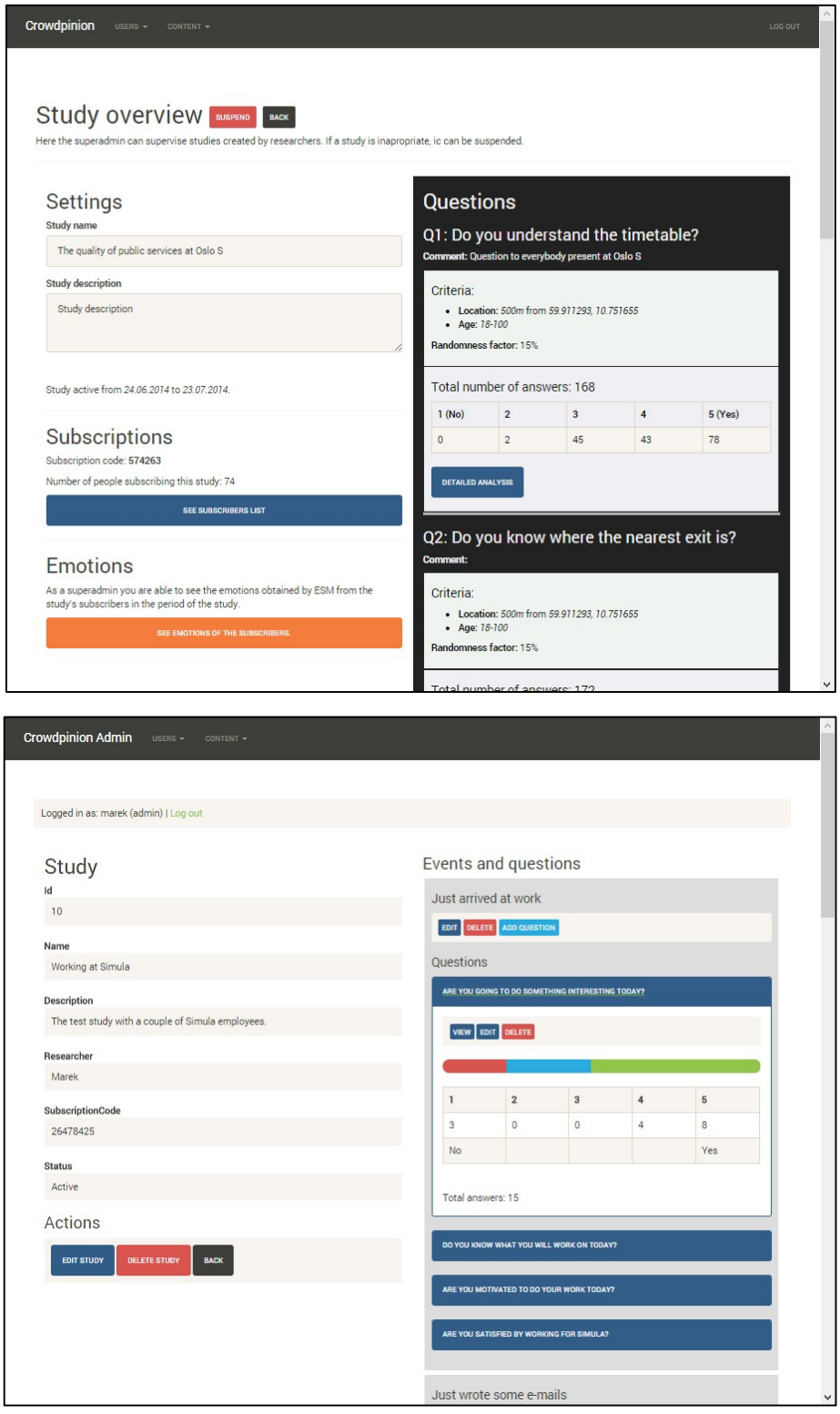

Figure 27. Comparison of mockup and real layout

In **Error! Reference source not found.** you can see an example of a mockup page and the final study view page. It is obvious that they differ quite a lot. I would say that the mockups have been more complex in terms of proposed functionality, but simpler in

styling. Some of the concepts from the mockups have changed, have been moved to the future work or dropped. Still the mockups have been very helpful when I have been discussing the solutions with my Supervisor, as they illustrated our ideas and made it easier to imagine the flow. I have also been able to use large parts of the layout of the mockups as the layout and styling of the current web panel in CakePHP.

Having the initial set of requirements, I started coding the web part of the system. The decision to start with the web panel was motivated by the fact that I have already been familiar with CakePHP and database design, so I could set the scaffolding of the web panel quickly and continue with its more complicated features and move to the mobile app.

I tried to work in an Agile way, deploying often (presenting the results of the work to my Supervisor every two or three weeks) and dynamically react to the changes of requirements. Unfortunately, the irregularity of some events at the university and in my professional life made it too hard to follow a regular Scrum routine of two-weeks-long iterations. Still the work has been going quite well and the entire Crowdpinion has been developed with only slight delay.

I have been testing the software continuously during the development. After developing a new feature and before committing the code to the SVN repository, I have performed a set of manual acceptance and regression tests of the new functions. Every few days I have been doing some more intensive testing of the entire system. Due to time limitations, I have not created any automated tests. I am aware that it is a very good practice to have them, but on the other hand, the size of the software in the current version makes it possible to test the entire system manually in reasonable time. If I go forward with developing Crowdpinion into a much bigger system, I will have to include a good set of unit test and automated regression tests.

There has been two major tests with the users. One of them has been conducted at Simula Research Laboratory and the other during the Second Workshop on Gamification in Information Retrieval GamifIR 2015. This has been a unique possibility to do beta testing of the software in the real environment of potential end-users. More information about these two cases can be found in the next chapter.

4. Evaluation and lessons learned

In order to evaluate Crowdpinion as an ESM surveying system, I have conducted three studies along the development process. The first study has been a small survey on a prototype done to evaluate the concept. The second and the third study were done with fully functional beta version of the software in two different contexts: work environment evaluation at Simula Research Laboratory (http://simula.no) and workshop evaluation at GamifIR' 2015.

4.1. Evaluation study I: Evaluating the concept

This study is a part of the paper that I published together with Michael Riegler and Sagar Sen at The Second International Workshop on Gamification for Information Retrieval (GamifIR'15) [12].

> In order to evaluate our approach to the method for obtaining information and the use gamification elements in Crowdpinion, we have conducted **a short survey with 10 participants**, using the early beta version of the app. We presented the app to nine students at the University of Oslo and one postdoctoral researcher at the University of Trento. None of the participants have been involved in or were aware about the project before. Two of the participants were female and eight were male. Six of them had a background in IT. In the evaluation, we described the purpose of the app, brief event-contingent ESM studies theory, the use case and made a quick walkthrough. After this each of the evaluation participants were given the following instructions:
>
> *Imagine you are asked to take part in this study for a month, every day when you are at work. Please answer the questions:*
>
> ***RQ2.1 Would you still be motivated to respond to the events' questions regularly?***
>
> ***RQ2.2 Would the ranking and the unlocking of the results help to keep you motivated?***
>
> ***RQ2.3 Are there any ways, which would work better for your motivation?***
>
> The general trend in the responses have been quite positive for our idea of adding gamification (the leaderboard and the unlockable elements) to Crowdpinion. Half of the responders stated that they would be motivated to take an active part in such study for a long period of time (**RQ2.1**). What is more significant, in the second question (**RQ2.2**) most of them (8 out of 10) said that the ranking and the unlocking of the summaries of responses would play a major role in building the motivation (Figure 28).

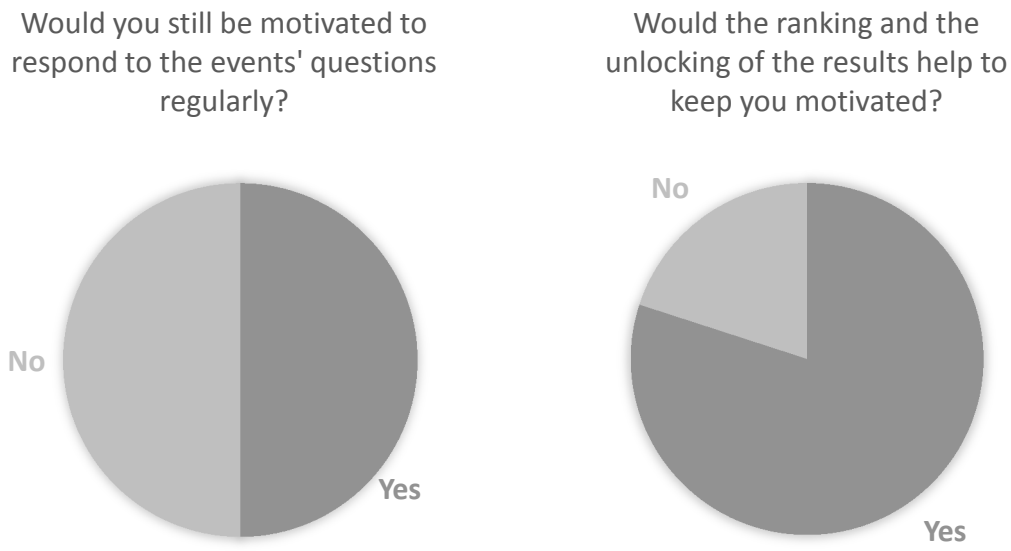

Figure 28. Responses to the first two questions.

There were several interesting inputs regarding motivation. One responder admitted that even if he would be motivated, he *might be forgetting to respond regularly*. Another person highlighted the *feeling of obligation to respond* after subscribing to a study as the key of her motivation. One responder stated that she is *not used to mobile apps* and therefore would not like to use Crowdpinion regularly. A couple of people mentioned that the *topic of a study must be interesting* to them if they are to be motivated and one said that he would be motivated, if he saw that *the study brings consequences* (*a real change*) in his environment.

Asked about what else could improve the participant's motivation (**RQ2.3**), one person emphasized the *technical quality of the app and it's interface* as a motivating factor, another said that (e.g. in a company context) it would be interesting to *divide responders into groups* (e.g. by teams or professions) and *compare responses between the groups*.

One of the responders suggested *extended gamification* – a game, where the participants answer questions while travelling through a maze. Each question would be located in a room with five closed doors and by choosing an answer, the participant would open one of the doors and go to the next room. At the end of a study, the responder would reach some point or escape the maze. During the game, the participants would also be able to see who else is following the same path – who is giving the same or similar answers. We find this idea interesting, because it would definitely improve the responders' immersion in the studies.

4.2. Evaluation study II: Work environment

As soon as Crowdpinion has reached a stable and fully functional level of development, I designed a study to test the system in a possible real application context. Evaluation of work environment has been mentioned before as one of the use cases and I selected it for the topic of the first study. This study consisted of two parts: first there was a momentary study using Crowdpinion and after the first part finished, I conducted a retrospective survey about the same topic and compared the results. As a part of the retrospective survey I have also asked the participants to evaluate Crowdpinion as a tool and they gave me valuable feedback about my work. In the subchapters of this study, I describe the set up and results of both studies, compare the results and conclude with the feedback I received about Crowdpinion.

In this evaluation study, I have been searching for answers to following research questions:

- **RQ3.1. What types of information can be extracted from results of a study in Crowdpinion?**
- **RQ3.2. Are the results from momentary study different from results from the retrospective study?**
- **RQ3.3. What motivates participants to taking part in studies?**
- **RQ3.4. How do the participants perceive studies in Crowdpinion?**

I decided to do the study at Simula Research Laboratory (https://simula.no). It was suitable for such study as a modern, medium-sized creative work environment. It was also convenient to do the study there, because I could easily contact a group of potential participants and persuade them to participate (because people at Simula are usually willing to help each other in the studies).

4.2.1. Momentary study

Study setup

The participants group consisted of nine full time employees. They were given a detailed instruction for the study and a manual for the mobile app. They were asked to actively participate in the study in at least 4 to 5 days at work. They participated by responding to sets of questions assigned to events that occur at Simula every day.

The scope and the size of the study has been set up in a way that would make it more or less similar to a real study and at the same time, I tried to keep the study relatively small. I was concerned that the participants would not agree to take part if it was too big and therefore too burdensome to participate. The study setup consisted of 9 events and 18 questions. The events, questions and possible answers can be seen in **Appendix A. Results of momentary study "Working at Simula"**.

Results and discussion

Because of the specific work schedule of the research employees, which involves travelling to conferences, the start and end dates vary for each participant. The first participant started on Monday 17.03.2015 and the last finished on Friday 27.03.2015 (Figure 29). The average participation period per participant was 3.44 days.

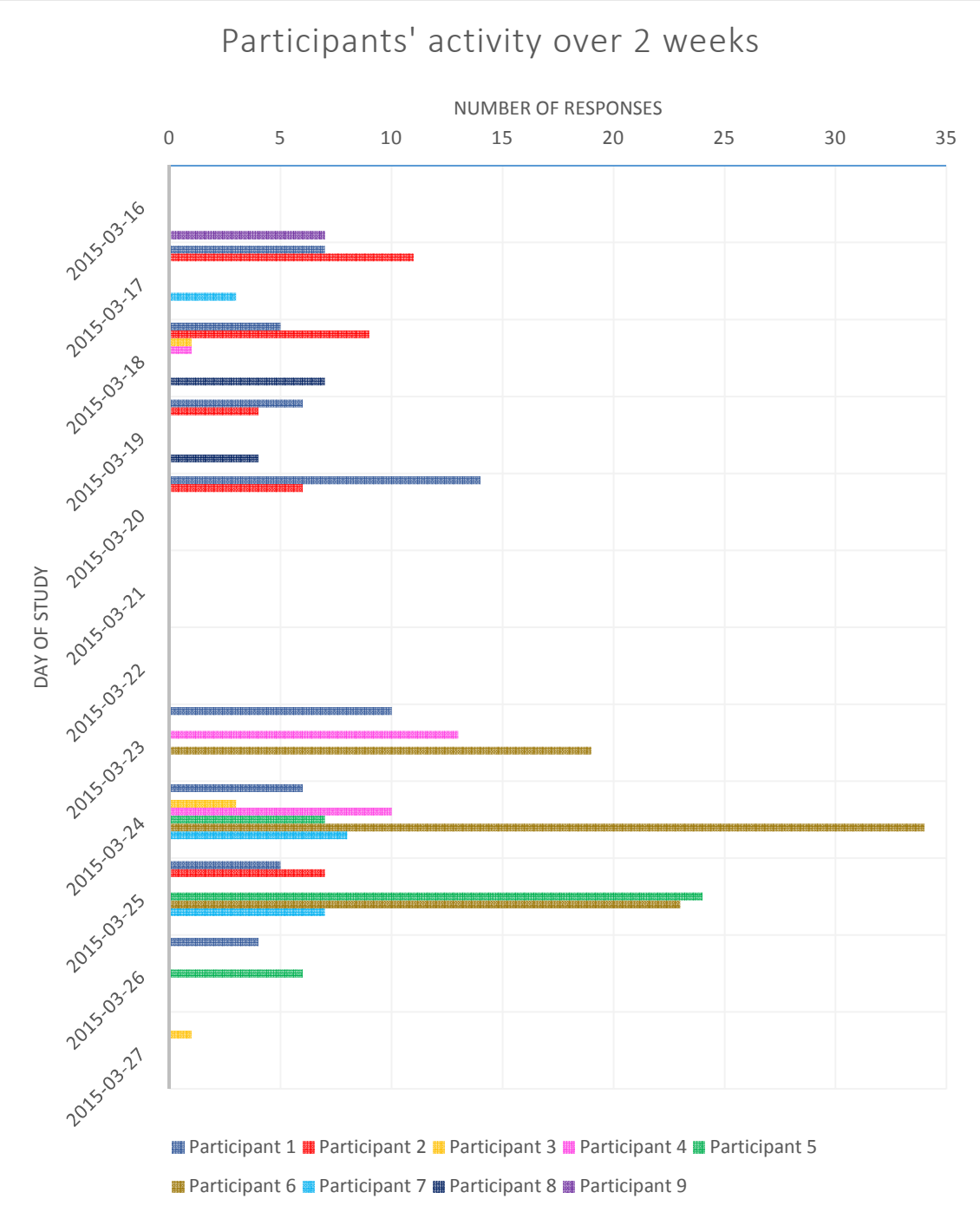

Figure 29. Participants' activity

There were several questions asked by the most active participants to this study (they are also presented in **Appendix A. Results of momentary study "Working at Simula"**). All of them have been relevant and did not require moderation. They did not get many responses though, because they have been asked quite late and the test study have been limited in time.

There are many facts that can be derived from the results of the study (**RQ3.1**). They can be analyzed in many ways:

- Independently as a set of summaries of responses to each question – to discover the overall trends.
- As set of answers with timestamps – to observe how the values have been changing in time.
- As couples, triples etc. of questions – some questions can be paired to see how similar or consequent questions are answered in relation to different events.

Analyzing a single question independently (Figure 30) it is easy to see that the participants usually know exactly what they are going to do at work, when they arrive to the office in the morning. This is a quite positive input on tasks management and planning.

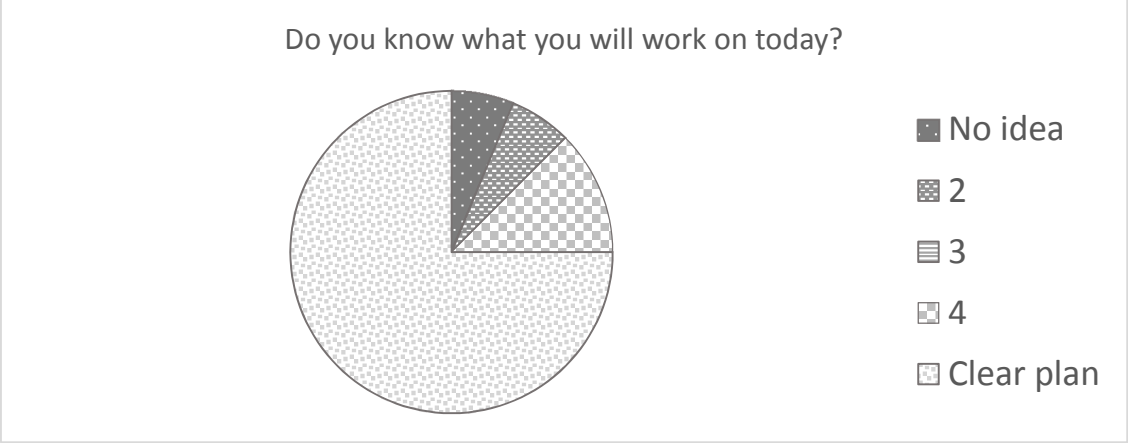

Figure 30. Do you know what you will work on today?

An interesting example here is the question about lunch preferences, where all responses clearly indicate that the participants prefer having lunch with their colleagues. The employees that used Crowdpinion have also indicated that they are usually happy to talk to their colleagues and at the same time, the conversations are rarely related to their work (Figure 31).

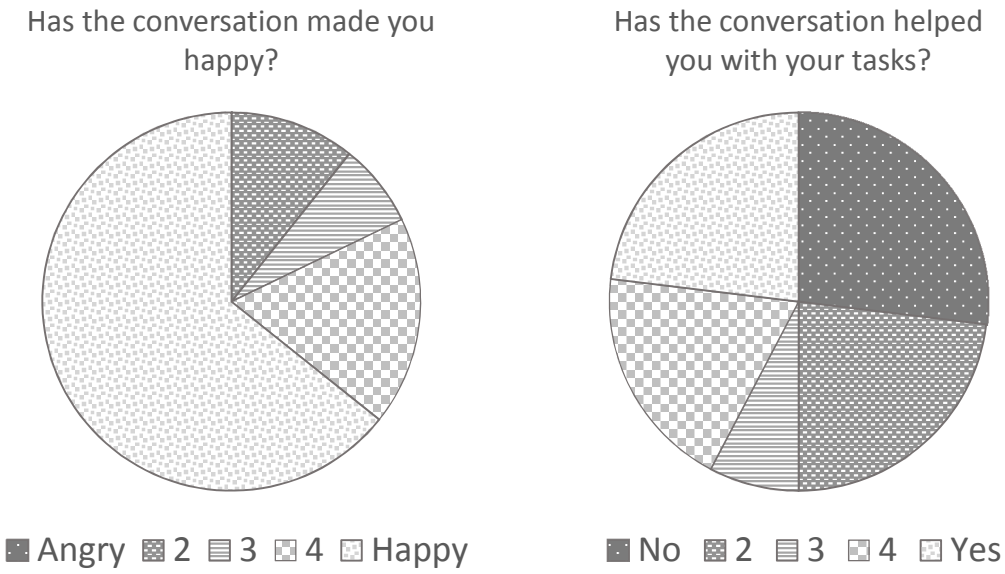

Figure 31. Opinions about conversations at work

In order to see interesting information based on timestamps, we need to choose questions that relate to events that can happen in many different times during a period of time. In this study, I was asking the participants about their motivation to work in the remaining hours, when they were having a coffee break. Coffee breaks at Simula do not have a fixed time and the timestamp of the responses varies from 9:53 to 21:03.

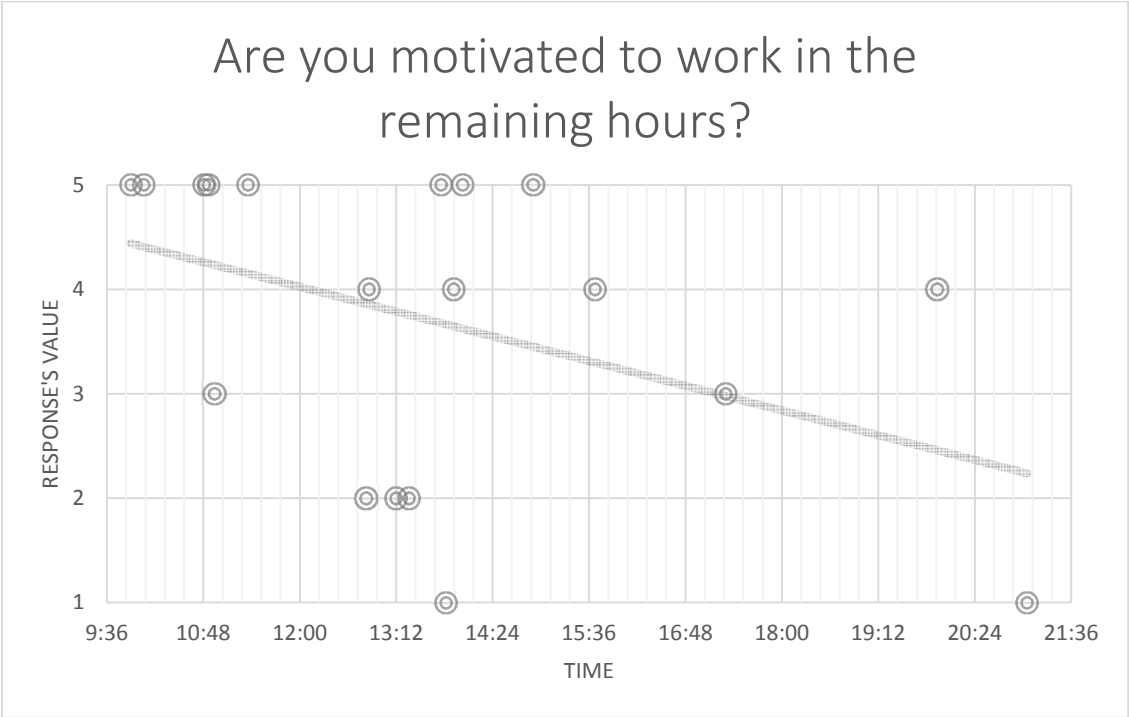

Figure 32. Motivation chart

Looking at the chart of the results, we can observe the times when the participants were having their coffee breaks and the trend line of motivation (Figure 32). It shows that the motivation decreases slightly during a day, with mostly high values in the morning, and lower values after lunch.

The last two responses have been given quite late - 19:55 and 21:03. It can mean that the employees stay in the office until late evening, which is a habit of some Simula employees. However, it can also indicate, that some participants were not responding to the questions in the right moments – perhaps they were using the app at home in the evening.

To verify if the participants really respond to questions immediately after the events, let us have a look at two questions - *Are you motivated to do your work today?* (Just arrived at work) and *Has it been a good day?* (Just before going home). These questions should be answered respectively in the morning and late afternoon.

Figure 33. Timing of responses

The results (presented in Figure 33) show clearly that it is not always the case. It seems that sometimes the participants respond to questions later than they should. This problem can be solved by using context-awareness to remind the participants about the events and to assess the quality of data that I describe in the future work section.

In this study, we can observe an example of interesting data that the researchers can solicit by asking the participants similar questions in different moments (events). In this case, the employees were asked a pair of questions:

- Are you going to do something interesting today? (when **just arrived at work**)
- Have you done something interesting today? (**just before going home**)

The responses to these two questions illustrate that the reality does not always match the expectations, but the employees are optimistic about their work (Figure 34).

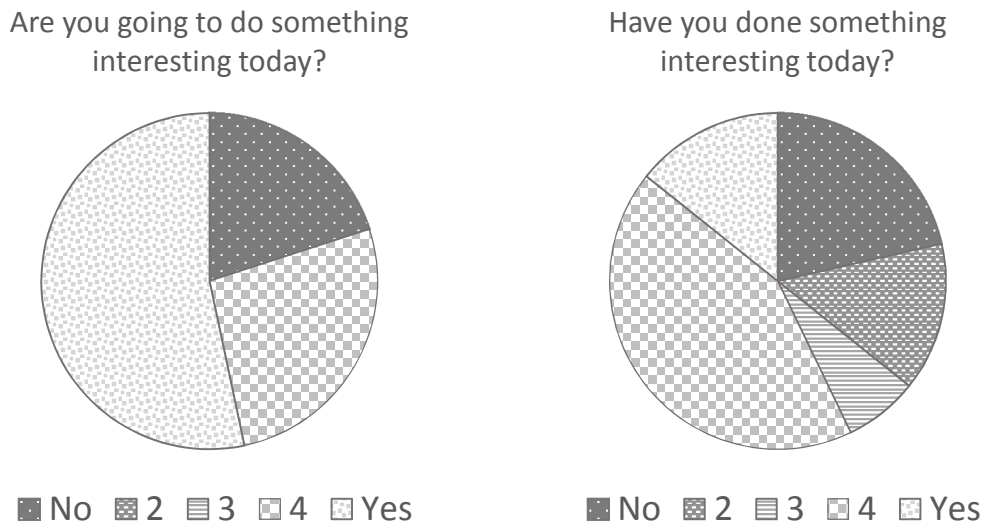

Figure 34. Expectations vs reality

4.2.2. Retrospective study

After two weeks of the momentary study, and a week of break, I have performed the second part of the study. The goal was to answer the research question **RQ3.2** by comparing the results from the retrospective study to the previous momentary study.

Study setup

I have created an online survey and in its first part I have asked the participants to respond to questions about some aspects of their work in March (it was the first week of April). It was then a typical retrospective survey when the participants had to think about situations that too place some time ago and evaluate those using single values – as opposed to evaluating the momentary situations multiple times in the previous part of the study.

The list of questions in this study included:

- Were you happy at work in March 2015?
- Did you have any interesting tasks at work?
- Were you satisfied with your work on your tasks?
- Have you been motivated to work?
- Were you often tired?
- Did you like writing e-mails?

- Were the lunches at work tasty?
- Were the meetings at work mostly productive?
- Did you like chatting with your colleagues?

The questions are clearly related to the questions from the experience sampling study. I deliberately obtained the opinions about the same issues from two types of studies so that we can have a look at the differences in the data.

Results and discussion

In the tables in **Appendix B. Results of the retrospective study compared with momentary study** and charts below, I present the comparison of responses to corresponding questions from both studies.

In order to answer **RQ3.2**, I would like to analyze the relation between responses to several questions. It is worth emphasizing that the both sets of responses were given by the same participants, so theoretically each participant should have the same opinion in both cases.

The number of responses in the momentary study is obviously larger than in the retrospective study – the participants were responding to the same questions many times. Therefore I have decided to present the ratio of responses in pairs of pie charts.

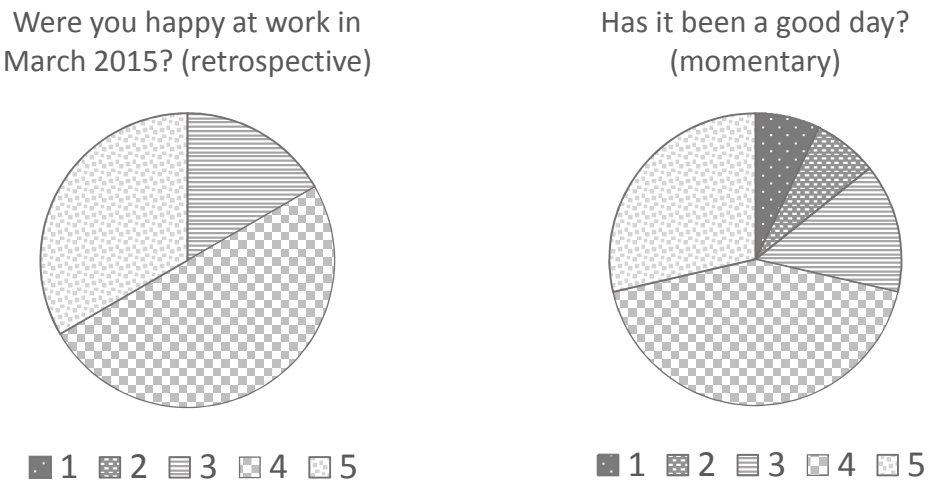

Figure 35. Comparison of responses about happiness

In some questions, the ratio of responses in both studies was quite similar. For example, in the comparison (Figure 35) of retrospective question “Were you happy at work in March 2015?” with corresponding momentary “Has it been a good day?”, the distribution of the responses is rather similar (take away the two single values of 1 and 2 in the momentary study).

It is not always like this though. In other questions there are remarkable differences between the sets of responses. In case of the pair of questions about interesting tasks

at work (Figure 36), the participants evaluated this factor more positively in the retrospective study, while they were not so satisfied when they were evaluating it momentarily at the end of each work day. This can also be compared with the momentary expectations versus reality comparison presented in Figure 34.

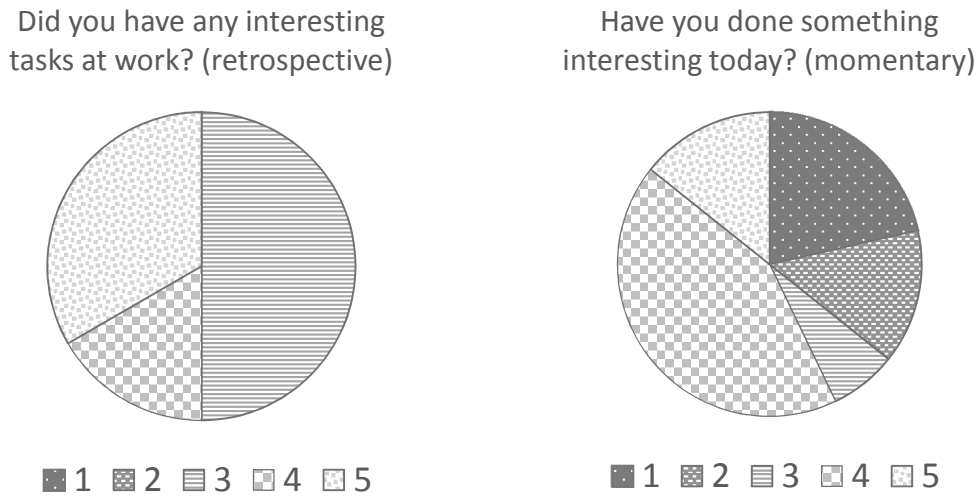

Figure 36. Comparison of responses about interesting work

In other cases, the responses from the retrospective set were more negative than in the ones from the momentary study. For example, in momentary reports just after a conversation with a colleague, the participants in most cases were very happy about it (Figure 37). The same participants were evaluating the happiness from meetings a bit lower in the retrospective study.

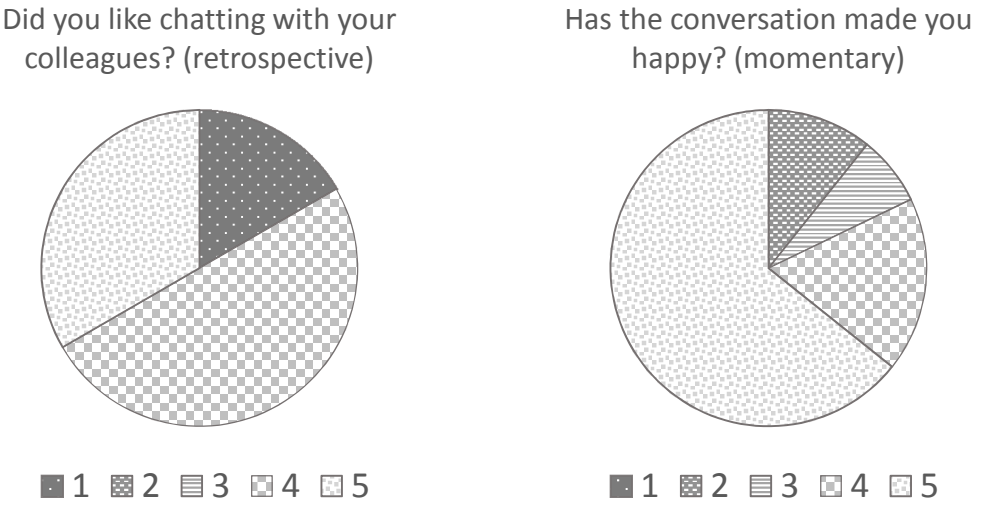

Figure 37. Comparison of opinions about happiness from conversations

In order to have a clearer picture and give the answer to **RQ3.2**, let us analyze the table with full set of average responses' values from all pairs of questions. The difference is the value of average retrospective – average momentary. If the difference is larger than zero, the average for the retrospective is larger than momentary.

Table 3. Average retrospective values vs. Average momentary values

Question topic	Average retrospective	Average momentary	Difference
Happiness at work	4.17	3.79	0.38
Interesting tasks	3.83	3.14	0.69
Satisfaction from tasks	3.67	3.20	0.47
Motivation to work	3.50	3.68	-0.18
Tiredness	3.83	3.83	0.00
Happiness from mailing	3.00	3.79	-0.79
Taste of lunch	3.67	3.25	0.42
Productivity of meetings	4.00	4.00	0.00
Happiness from conversation	3.83	4.36	-0.52

The values in Table 3 show that for most questions there is a noticeable difference between the average of responses in the retrospective and in the momentary study. In four questions the participants were more positive in the retrospective study. In three questions they gave higher responses in the momentary study. The average in two questions has been exactly the same for both studies.

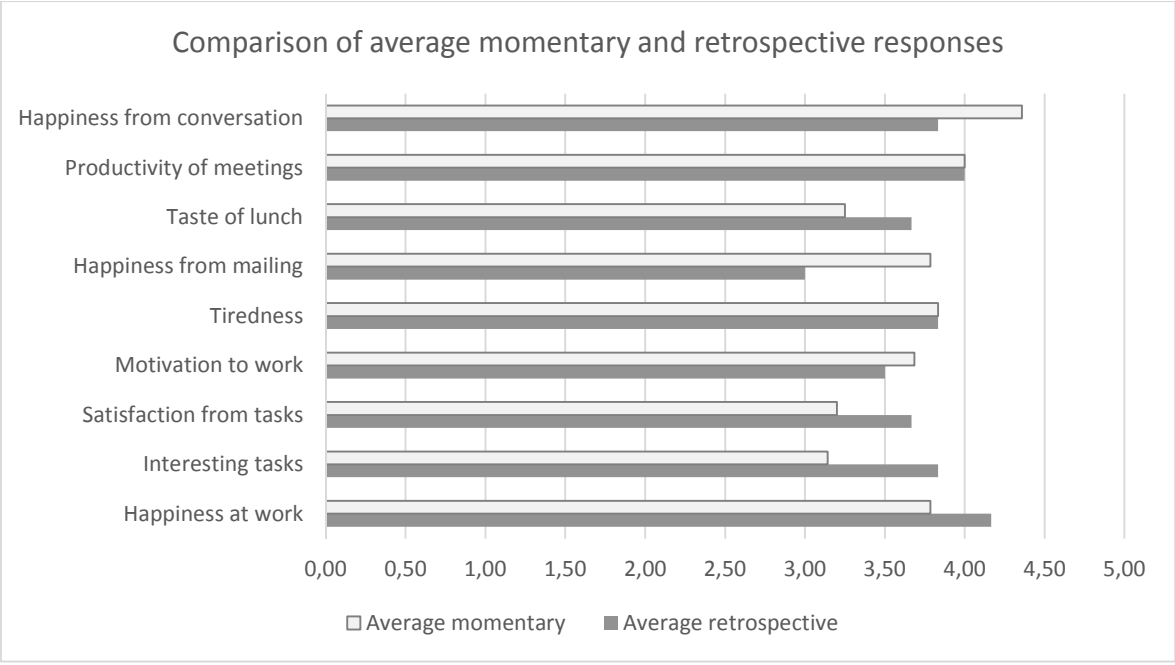

Figure 38. Comparison of momentary and retrospective responses

It can therefore be observed that the participants tend to respond differently in retrospective and momentary studies (Figure 38). The participants in retrospective studies can perhaps be biased by the current situation or by memory – it is possible that they remember just some fragments of the past.

4.2.3. Evaluation of Crowdpinion

Study setup

In the second part of the retrospective survey, I have asked the participants to directly evaluate Crowdpinion as the tool that they were using in the momentary study. I have prepared a set of questions that were inquiring about some areas of the tool that I particularly needed feedback or information about. Inquiring about answers to **RQ3.4**, at the same time I have been asking for feedback for my future work. A question about motivation attached to this part of the study allowed me to gather information needed to respond to **RQ3.3**. The list of questions and summaries of responses can be found in **Appendix C. Evaluation of Crowdpinion**.

Results and discussion

The participants responded positively or moderately positively to most of the questions. The part that made me happy in a special way is the decisive five for the question about quick and easy responding workflow. This means that I have managed to fulfill the main design requirement of the app. The participants liked the design of the app and mostly enjoyed participating in the study.

Some people indicated that the study (meaning the set of events and questions) was not designed well. I acknowledge it, as it has been my first study and there might have been some flaws in the study setup. The responses to the last question indicate that the participants use the gamification elements a bit less often than expected. The ideas on how to apply gamification effectively can be reconsidered.

A subsection of the survey was about evaluation of each gamification element separately. The elements have been rated quite high in terms of improving engagement and motivation (Table 7). The leaderboard and the unlockable summaries of responses got the average of 3.83. The third element, the unlockable possibility to add participants' own questions have been rated little lower (3.50). This can be caused by the fact (emphasized by one of the participants) that the threshold of 50 responses that allowed the participants of this study to unlock this feature was set too high. The participants were unlocking it late or not at all. This indicates that in the future versions of Crowdpinion I should add the possibility to configure the thresholds, so that they can be lowered if they prove to be set too high.

In the survey, I have also included two larger multiple-choice questions. The first one is addressing the third research question (**RQ3.3**) and the other is strictly about Crowdpinion (**RQ3.4**). The participants could select as many answers as they wanted.

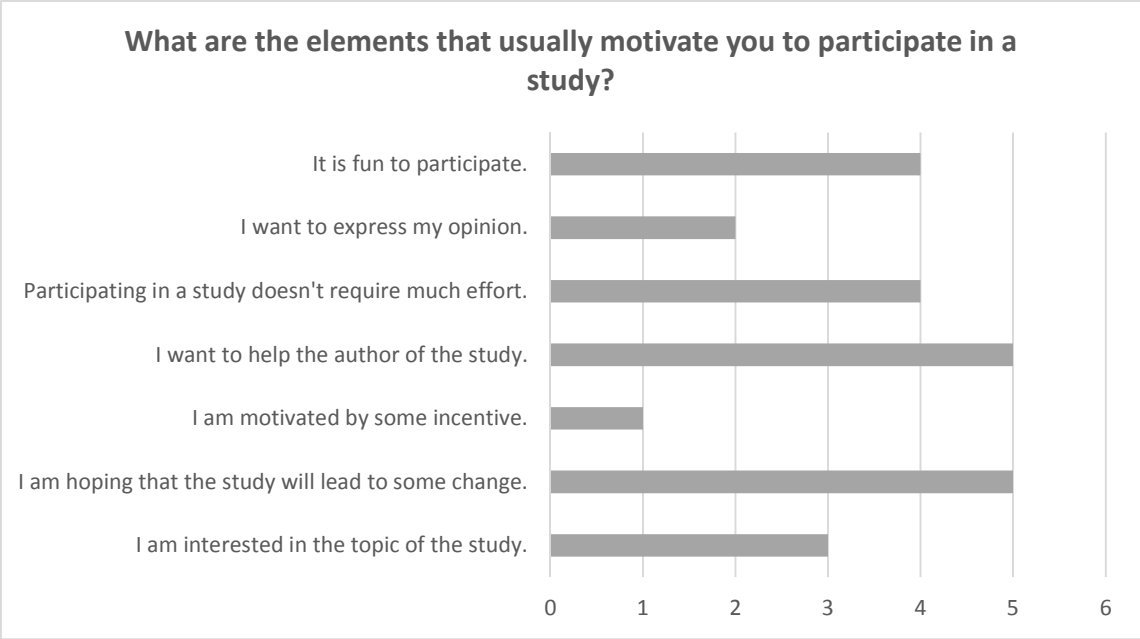

Figure 39. What motivates the study participants?

As a response to **RQ3.3**, the participants gave me a good insight into the factors that usually motivate them to participate in a study (Figure 39). The options that were selected by most people were:

- I want to help the author of the study.
- I am hoping that the study will lead to some change.
- It is fun to participate.
- Participating in a study does not require much effort.

Especially the last two points are important for me as the developer of Crowdpinion. The participants expect the studies to be effortless and enjoyable. I believe that Crowdpinion already matches these requirements to some extent, but the results also indicate a possible direction of the further development.

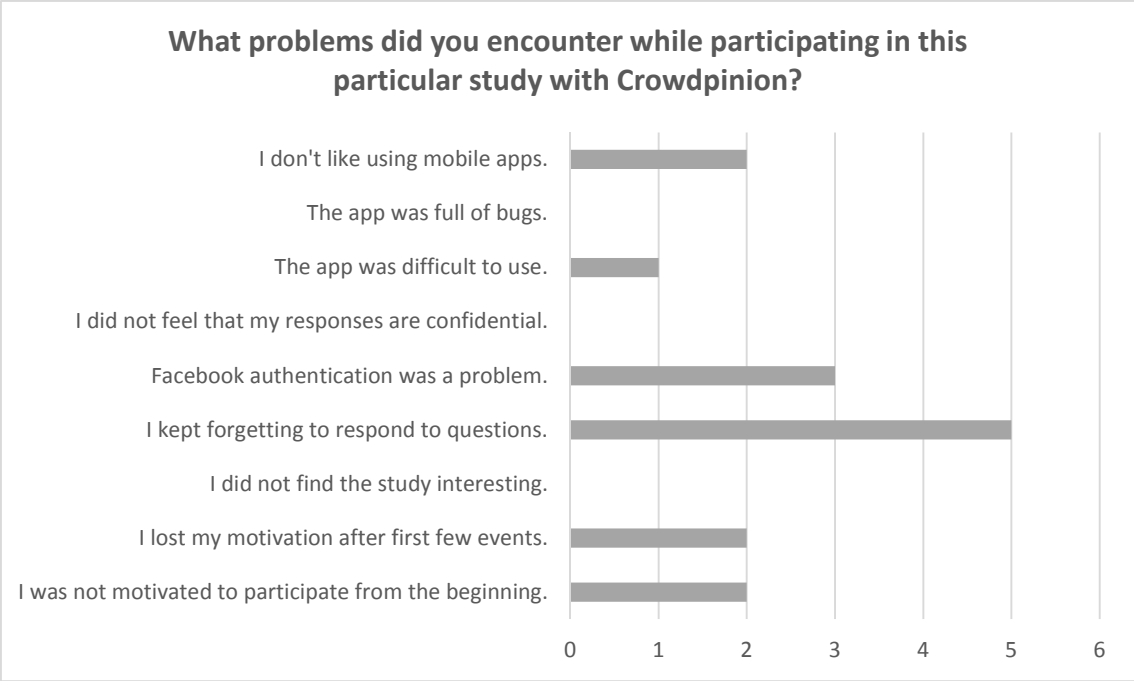

Figure 40. What problems did the participants encounter in this study?

When asked about the problems that they encountered in the study and the app (Figure 40), the participants selected two main issues:

- Facebook authentication was a problem.
- I kept forgetting to respond to questions.

These two results are important issues that should be fixed soon. I have been getting suggestions about these two factors from the participants for some time. What I expected to be an easy solution to the authentication issue, became a factor that is undermining the feeling of anonymity – people do not feel anonymous when they log in with Facebook, even if only the minimal amount of information is actually retrieved from their profiles and they are anonymous to the researchers.

The issue of participants being motivated to participate, but still forgetting to respond has also appeared earlier in the study. It is quite likely to happen, when it is the participant who triggers the response in the event-contingent ESM protocol. If the situation will continue to be a problem, I will consider implementing at least some elements of the signal-contingent ESM, where the participants receive a signal (a beep, a vibration) at the moments, when they are required to respond.

4.3. Evaluation study III: GamifIR 2015

The paper that I wrote together with Michael Riegler and Sagar Sen was accepted to the GamifIR workshop at the 37th European Conference on Image Retrieval. We thought that it was a very good occasion to conduct another study with Crowdpinion. This study would be different from the one that I have had conducted before at Simula, because it would be only one day long and some events would occur only once.

Study setup

The workshop's chairs welcomed our idea with enthusiasm and I prepared the study. I was hoping to test the system in a big group of participant. However, apparently only around 15 participants registered for the workshop. Out of these, less than half had Android devices and some people were encountering some problems with the Internet connection, which made it difficult for them to react to every event. Eventually there were six participants.

By conducting the study, I wanted to test Crowdpinion in context of a short-term study, therefore the only research question in this evaluation study is:

- **RQ4.1. Can Crowdpinion be used in a short-term study?**

Results and discussion

Despite the low number of participants, I still think that the study was very interesting and worth the effort. The participants gave 145 responses. It gives the average of almost 25 responses per participant, which is a good result in a study that took only around eight hours.

The questions and responses from this study can be found in **Appendix D. Results of study at GamifIR 2015**.

The results of the study have been presented to the audience during the last presentation. The general trend in the results have been positive and it provided the participants and the presenters with some interesting information. Several examples of the responses summaries can be found below.

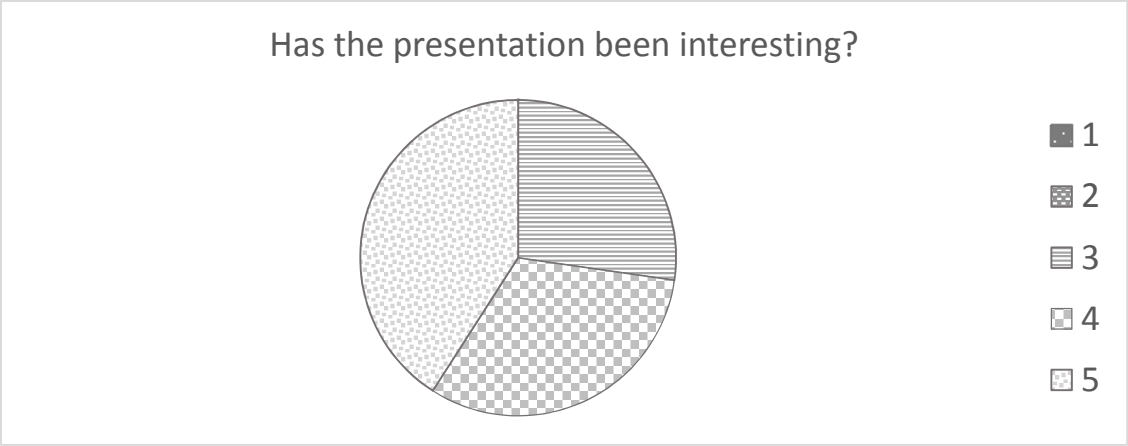

Figure 41. Has the presentation been interesting?

Happily for the presenters, the study participants rated the presentations as interesting (Figure 41). Nobody selected the two lowest values of responses.

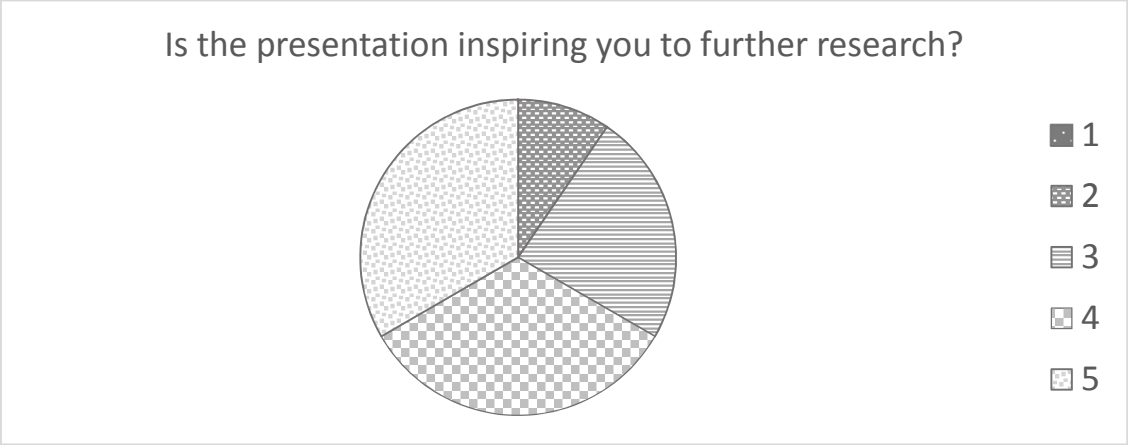

Figure 42. Is the presentation inspiring you to further research?

The participants have also found most of the presentations inspiring (Figure 42).

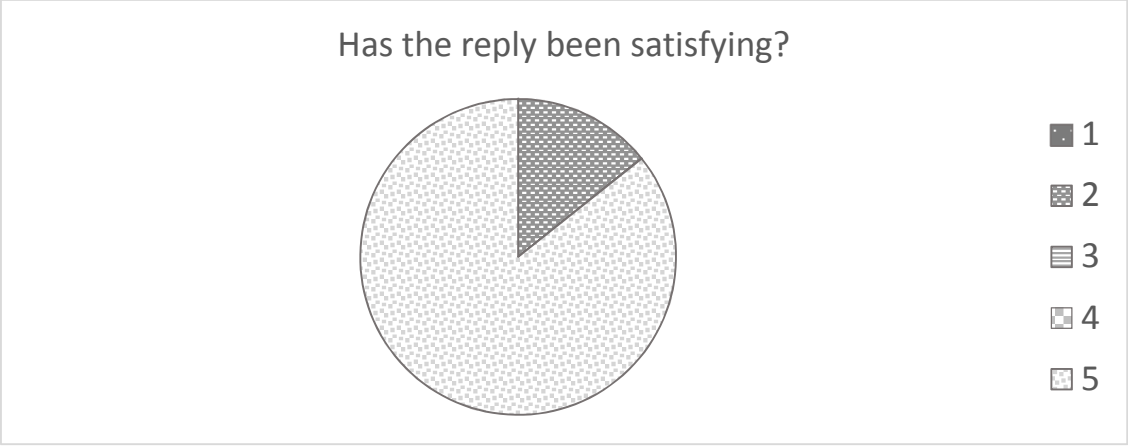

Figure 43. Has the reply been satisfying?

Most of the people who asked a question to the presenter have found the response satisfying (Figure 43).

In terms of the questions regarding presentations, I have decided to have a single "Listening to a presentation" event instead of separate events for each presentation. It was partly because of discretion – knowing that I was most likely going to display the results of the study at the end of the workshop, I did not want to make the presenters embarrassed in public if the evaluation of their presentations were particularly low.

The results have been rather positive in the end, so nobody would perhaps be upset with the rating of their presentation. Slight differences can be observed in the chart of average values of three factors from the presentation-related questions: Focus, Inspiration and Interest over time (Figure 44). The average values are lower in the second block of presentations. Perhaps the quality of the presentations was a bit lower. It is also possible that study participants were tired and perceived the presentations as less interesting.

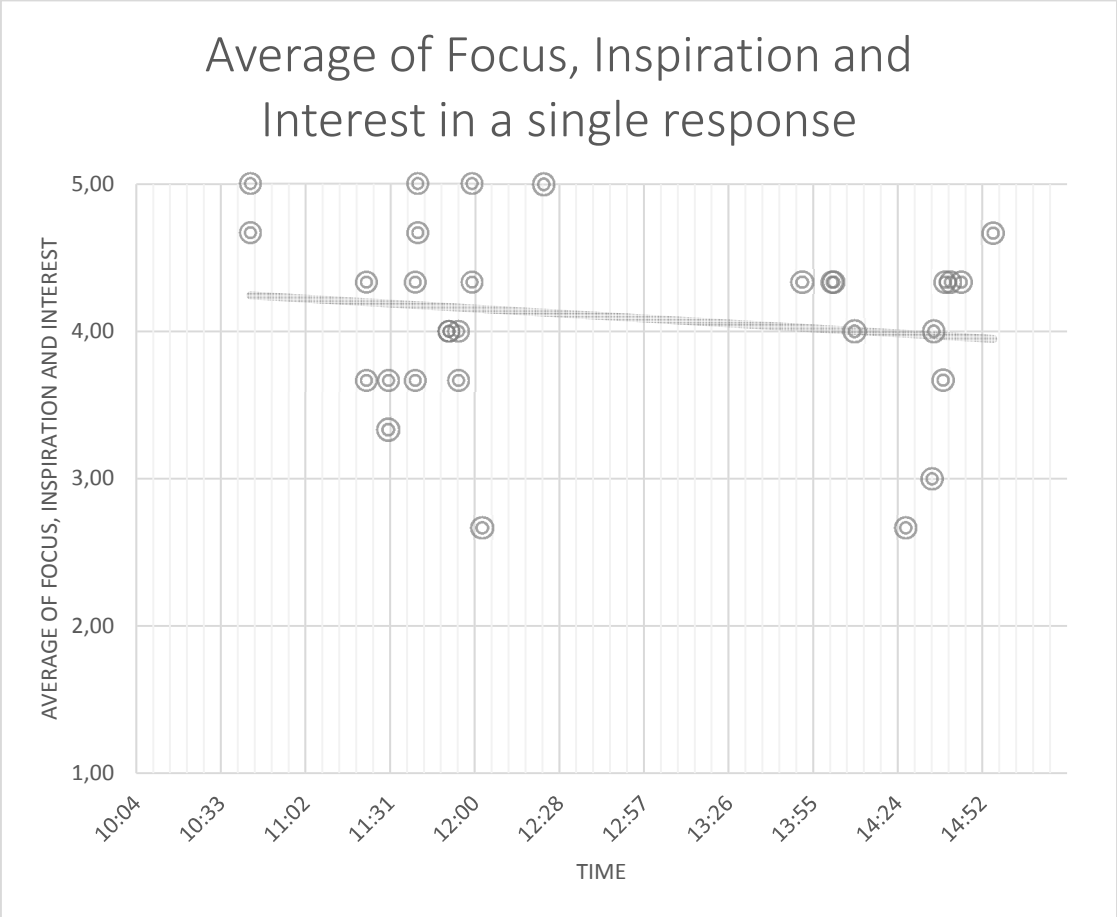

Figure 44. Opinions about presentations

As an answer to **RQ4.1**, I can say that Crowdpinion can be used in short-term studies. The participants use it actively, even though they do not make much use of the gamification elements. Responding to questions did not interrupt the participants in the activities at the workshop.

After the study, I have talked to the people who used the app. Even though there were so few participants, I have received a decent amount of feedback. Two people pointed out that the login with Facebook can cause several problems. One person said she did not remember her Facebook password, because usually when she used Facebook, the password is stored in the device or browser. Two participants emphasized that for them using Facebook for authentication damages the feeling of anonymity and without anonymity they do not feel comfortable to answer some controversial questions.

Another issue that has been emphasized by the participants in this study is that while being immersed in the presentations and other activities, they were forgetting to use the app. This can potentially lead to unwanted situations when the participants respond retrospectively long after the events of interest or not respond at all. One person suggested that the signal-contingent protocol of the Experience Sampling method [16] where the device reminds the participant to respond, would make it easier to respond in the right moments. Introducing content-awareness that I describe in the future work chapter, can also be a solution to this problem.

5. Summary and future work

5.1. Future work

I am aware that Crowdpinion is still far from being a perfect momentary opinion surveying system. I have collected a list of features that I would like to include in further development of the system. The feedback that I have received from people who used Crowdpinion in the evaluation studies also several good ideas too. This section contains several ideas from my previously published paper [12] and some more that appeared later in the development.

5.1.1. Security and Anonymity

Even if the amount of data and sensitive information is rather small in Crowdpinion, the system must be well protected in terms of security. The authentication and authorization mechanisms provided by CakePHP protect the data. Still before the final release of the system I would like to perform a thorough security analysis and make sure that there are no technical security issues.

Apart from the technical part, there is another aspect of security, which is perhaps even more important. Anonymity in Crowdpinion is an issue on the conceptual level as well. This means that even if all security measures are done well technically, there is a risk that the participants can still be identified. It is because of the responses – in some situations the researchers can identify the participants by when and how they respond to questions. Let us imagine a study about working environment. If there is a day when there is just one meeting with four participants, and after this day there are four new responses to a question about productivity of meetings rating it very low, the researchers can with high probability assume that these four specific participants of a meeting have negative opinion about productivity at work. If information about this reaches the management, the four employees can face unpleasant consequences.

This issue was emphasized by participants in the study at Simula and at GamifIR. In both cases, the participants found the Facebook login mechanism particularly inconvenient. They said that having to log in with their personal Facebook accounts does not give them the feeling of anonymity. They also said that when they do not feel anonymous, they will hesitate to give negative answers in controversial questions. It is an important lesson and as one of the first changes that I will do in the future development of Crowdpinion will be replacing the Facebook login with a custom, more anonymous solution.

5.1.2. Social Targeting

An idea which I find quite promising in terms of making Crowdpinion a better research tool is related to social targeting. In the future I would like the app to encourage the users to provide some more personal details (social parameters) about themselves, that would allow classifying the app users by age, location, gender, education, profession etc. Having this classification, Crowdpinion would enable the researchers to make studies in which they would send questions only to selected categories of participants and analyze the results in regard to the social parameters later. For example, conducting a study about awareness of risks of breast cancer, the researchers could specify that the questions in the study should be accessible only by women between 15 and 60 years old, living in Oslo. The app could then suggest suitable participants that the researchers could invite. After the study, the researchers could filter the results by different criteria and for example search for correlation between the awareness level and education.

5.1.3. Context-Awareness

Crowdpinion is used on smartphones – devices which are equipped with a very functional set of various sensors: GPS, clock, accelerometer etc. Furthermore, the operating systems at the devices operate on data of its users, for example store calendar entries. If Crowdpinion is developed further so that it is able to access and analyze the data from the sensors, it can make a great use of context-awareness [104]. An example of application of context-awareness is Google Now, which can guess the context of the users very accurately based on the data from the sensors.

Crowdpinion could use context-awareness to predict the events that can be happening to the users at the moment. For example if it detects that the user is approaching the Oslo S train station, it can notify the user and suggest the event “Going to take a train” that is a part of a study about public transportation in Oslo. If the app had access to the calendar, it could learn about times of scheduled meetings and ask the user the related questions just before and just after each meeting. As pointed by one of the participants of the study at GamifIR, these notifications would probably largely improve the number of events to which the users react.

5.1.4. Quality of Data

The data about users and their contexts can also be used for another purpose. Having this data, I could attempt to create an algorithm for evaluating the value of quality of data [105]. Quality of data is a property showing how likely it is that the particular response is true for the participant and given in the right context. For example, if there is a set of questions related to a “Just arrived in the office” and an user responds to the questions while the information from the phone sensors tells us that it is Saturday evening and the participant is in a pub in central Oslo, the value of the quality factor is

low. On the other hand, if the participants responds to questions about “Outdoor free time activity” while being in the mountains on a sunny Sunday afternoon, the value will be high. If all responses have a quality value, the researchers can decide that they want only responses with the quality factor higher than some threshold (e.g. 75%) included in their studies’ results.

5.1.5. Extended Gamification

Gamification has been an important element of Crowdpinion and the evaluation studies showed that it does have a positive influence on motivation and engagement of participants. It is therefore worth spending more effort on implementing some extensions and improvements.

Crowdpinion could for example become a more social app where participants would interact with each other directly in some sort of multiplayer game. The maze that has been suggested by one of the participants of the first evaluation of Crowdpinion is one of the possible directions. In such maze game, the participants would be moving their avatar across a maze by responding to questions. Each segment of the maze can have up to five possible exits linked to five possible answer values of a question. While travelling through the maze, the participants would be meeting avatars of other participants who reached the meeting point by responding to sets of questions in similar way.

Going towards more advanced game design and more advanced technologies, Crowdpinion could largely benefit from augmented reality. Combined with context-awareness and the camera and display of smartphones, Crowdpinion could display questions in the surrounding of participants, stimulating them to actively search for opportunities to contribute to studies.

The gamification enhancements can have improve participants’ engagement, but can largely change the core functionality of Crowdpinion and damage its simplicity, which, according to the users, has been one of the biggest advantages of the app. They would also require very big development efforts and therefore need to be reconsidered.

5.2. Summary

In this thesis, I have described the work that I have done during my master's degree studies. During this time, I have been concerned about optimal ways of obtaining information from software users with a special focus on long-term studies on momentary opinion, meaning the short-term opinion that people have at the very moment of asking the questions. I chose the event-contingent protocol of the Experience Sampling Method (ESM) [15] as an efficient methodology for this kind of studies and implemented it in a surveying system that I called Crowdpinion.

I do believe that studies on momentary opinion can produce very valuable data, which can be combined with results from surveys-based retrospective studies or used separately. Having the data from a momentary event-contingent Experience Sampling [16] researchers can learn much more than from retrospective studies. For example, in a classic survey they can ask a question like "Did you feel tired at work in last weeks?" and obtain responses about what the participants *remember* about their *overall* tiredness at work. In an ESM study, the researchers can ask the same question "Are you tired now?" many times in different situations over long time. By doing this, they obtain much more data, which allows them to analyze how the feeling of tiredness changed in different times of a day, in different days and in relation to different events. It is also important that the responses in momentary studies are not biased by memory, because they are reported almost immediately after analyzed events.

Crowdpinion is a software system supporting event-contingent ESM studies. It consists of a web panel for researchers, a mobile app for studies' participants and server-side services that connect these two. The researchers create studies. Each study has a set of events. Each event has one or more questions. After setting up the study, the researchers invite the participants to the study by distributing a subscription code. The participants input the code in their app and can immediately begin to participate in the study. They select an event from the list whenever it happens and respond to a set of questions assigned to it. The responses are immediately sent to the database and the researchers can access them at any time in the web panel.

An important part of the app is the gamification [11] elements – elements of games that are included in the app to improve engagement of participants in long-term studies. I have included them because there was a high risk that in ESM studies that can take many weeks, only a couple of the most motivated participants would remain active after first few days. Crowdpinion includes features that can be unlocked by participants when they respond to questions in studies. These features make Crowdpinion quite a special surveying system. First, the participants can access summaries of responses to questions – at the beginning they see the summary for the first question and an information that the results of the second question can be unlocked by giving some more responses. When the participants respond to a big number of questions, they can unlock the possibility to add their own questions to the study. There is also a ranking of the most active participants, so that they can compare themselves to the others and challenge

each other. I have described the functions, architecture and implementation of Crowdpinion in **Chapter 3**.

In the theoretical part of this work (**Chapter 2**), I have analyzed pieces of existing literature about obtaining information from users, the experience sampling method, momentary user experience (UX) evaluation [9], crowdsourcing [23] and gamification [11]. The Experience Sampling Method and evidence about ways of asking questions and soliciting responses laid the foundation for the methodology of studies in Crowdpinion and the design of the system. The Systematic Literature Review [7] of momentary UX evaluation method added a valuable overview of the state of art in the field of one of the most promising possible applications of Crowdpinion and provided me with many ideas. The research in crowdsourcing has been included, because garnering information from users in long-term studies resembles crowdsourced generation of data and faces the same challenges. One of the challenges is keeping people motivated and engaged. This is why I have researched the field of gamification and implemented some of its elements.

In **Chapter 4**, I present the evaluation studies that I have performed on Crowdpinion. The first evaluation has been done when Crowdpinion was still a prototype and I had just added the gamification elements. I wanted to have feedback about the concept of the system and gamification. I have asked ten fellow students to make themselves familiar with the prototype and respond whether they would be motivated to use it in a long-term study and whether the gamification elements would have increased their motivation. Only half of them said that they would be motivated, but 80% of the participants found gamification motivating. I have also received plenty of input about possible features and improvements, which was very valuable at that stage [12].

The second and the third evaluation study has been done with a fully functional beta version of the system – the web panel deployed on a server and the app built for Android. I wanted to test Crowdpinion in a context of real studies. In the second study, I asked nine employees of Simula Research Laboratory to take part in an experience sampling study, in which they would actually evaluate some aspects of working environment at Simula using the Crowdpinion mobile app. Participants have been responding to questions related to various events happening at work – e.g. *"Do you know what you will work on today?"* at *"Just arrived at work"* or *"Has the meeting been productive?" "Just after a meeting"*. Each participant was asked to use the app at least four to five days at work, but some used it longer.

After the momentary study, I have conducted a short retrospective study with the same participants. I have created a survey with questions about the same issues (motivation, productivity, happiness etc.) that were evaluated in the momentary study. The comparison of how the same people respond to similar questions in momentary and retrospective studies showed interestingly big differences and supported the hypothesis that data from momentary studies is a valuable complement to data from retrospective studies. In the last part of the second study, the participants gave feedback about the Crowdpinion and insight about what drives them to participate in studies in general.

I conducted the last study at GamifIR workshop at the ECIR conference in Vienna. As a use case, this study has been different from the one at Simula, because it took only one day and some questions were being answered only once. Apart from rather small number of participants (6 people), Crowdpinion worked well and gave some insight about how people perceived some aspects of the workshop. I have also received some feedback about the system from the participants.

I have summarized the feedback from the evaluation studies and several ideas of my own in the *Future work* section in **Chapter 5**. I can improve and extend Crowdpinion in many ways. I could make it a better tool for the researchers by adding more possibilities to configure studies and analyze results. This can be achieved for example by using the information from smartphones sensors to learn about context and quality of data or allowing the researchers to target questions to particular groups of users. On the other hand, it could also be made a better app for the participants. User experience of the mobile app can be enhanced by better design, improved technical quality (e.g. in terms of security) and improving the fun coming from gamification.

Working on Crowdpinion and learning about practical applications of crowdsourcing, gamification, experience sampling etc. has been a very good experience for me. I will probably continue working on it in my spare time after graduation.

Glossary

Experience Sampling Method (ESM) – a research methodology created by by Larson and Csikszentmihalyi [15], designed to obtain participants' momentary opinion in unsupervised studies.

Momentary opinion – opinion that a person has at this moment and that can change, fade or disappear in a short time.

Crowdsourcing – process of obtaining data or solving problems by splitting it into small tasks and having the tasks solved by an usually large group of anonymous workers.

Gamification – use of game design elements in non-game context [11] in order to increase users' motivation and engagement in some action.

User Experience (UX) – people's emotions, feelings and experiences coming from use of a product.

Appendix A. Results of momentary study *"Working at Simula"*

Just arrived at work

Are you going to do something interesting today?

3	0	0	4	8
No				Yes

Do you know what you will work on today?

1	1	0	2	12
No idea				Clear plan

Are you motivated to do your work today?

1	2	2	4	7
Demotivated				Motivated

Just wrote some e-mails

Has the mailing made you happy?

0	3	2	4	5
Angry				Happy

Coffee break

Do you feel more focused after the coffee break?

1	4	3	4	6
Distracted				Focused

Are you motivated to work in the remaining hours?

2	3	2	4	8
Demotivated				Motivated

Lunch break

Would you rather have your lunch…?

0	0	0	0	14
Alone				With colleagues

Has the lunch been tasty?

2	4	1	6	3
Awful				Delicious

Just before a meeting

Do you expect the meeting to be productive?

1	0	0	0	3
Waste of time				Productive

Do you know the plan of the meeting?

1	0	0	0	3
No idea				Detailed plan

Just after a meeting

Has the meeting been productive?

2	0	0	5	6
Waste of time				Productive

Are you still motivated to work after the meeting?

0	3	1	2	7
Demotivated				Motivated

Just talked to a colleague

Has the conversation made you happy?

0	3	2	5	18
Angry				Happy

Has the conversation helped you with your tasks?

7	6	2	5	6
No				Yes

Just completed a big task

Are you satisfied with your work on the task?

1	1	0	2	1
Frustrated				Satisfied

Are you tired?

3	1	1	0	1
Tired				Full of energy

Just before going home

Has it been a good day?

1	1	2	6	4
Terrible				Excellent

Have you done something interesting today?

3	2	1	6	2
Boring				Interesting

Questions asked by participants

Are you satisfied by working for Simula? (in Just arrived at work)

0	0	0	0	4
No, I am not.				Yes, I am.

Did lunch discussion lead to better collaboration? (in Lunch break)

2	1	0	0	0
Just small talk				We are going to work on something new

Was the talk driving you towards a common vision? (in Just talked to a colleague)

1	1	1	3	0
No common vision				The vision makes work effortless

Do you feel overworked? (in Just arrived at work)

2	1	0	0	1
No				Yes

Table 4. Responses to the study at Simula

Appendix B. Results of the retrospective study compared with momentary study

Were you happy at work in March 2015? and
Has it been a good day?

0	0	1	3	2
Upset				Happy
1	1	2	6	4
Terrible				Excellent

Did you have any interesting tasks at work? and
Have you done something interesting today?

0	0	3	1	2
Not at all				Only interesting tasks
3	2	1	6	2
Boring				Interesting

Were you satisfied with your work on your tasks? and
Are you satisfied with your work on the task?

0	2	0	2	2
Frustrated				Satisfied
1	1	0	2	1
Frustrated				Satisfied

Have you been motivated to work? and
Are you motivated to work in the remaining hours?

0	1	2	2	1
Demotivated				Motivated
2	3	2	4	8
Demotivated				Motivated

Were you often tired? and
Are you tired?

0	0	3	1	2
Hardly ever				All the time
1	0	1	1	3
Full of energy				Tired

Did you like writing e-mails? and
Has the mailing made you happy?

1	1	1	3	0
No				Yes
0	3	2	4	5
Angry				Happy

Were the lunches at work tasty? and
Has it been a good day?

0	0	3	2	1
Awful				Delicious
2	4	1	6	3
Awful				Delicious

Were the meetings at work mostly productive? and
Has the meeting been productive?

0	0	1	4	1
Waste of time				Productive
2	0	0	5	6
Waste of time				Productive

Did you like chatting with your colleagues? and
Has the conversation made you happy?

1	0	0	3	2
Not at all				Very much
0	3	2	5	18
Angry				Happy

Table 5. Comparison of responses from the retrospective and the momentary study

Appendix C. Evaluation of Crowdpinion

Did you enjoy participating in the study?

0	0	2	1	3
Not at all				Very much

Was the app designed well?

0	0	2	2	2
Not so much				Well

Was it quick and easy to respond to the questions?

0	0	0	1	5
No				Yes

Was the study designed well?

0	2	0	2	2
Not so much				Well

Were you motivated to respond to the questions?

0	0	0	4	2
No				Yes

Were you interested in the outcome of the study?

0	0	1	0	5
No				Yes

Did you often use the gamification elements?

0	0	2	3	1
Never				Often

Table 6. Evaluation of Crowdpinion and the study at Simula

Did the leaderboard improve your engagement and motivation?

1	0	1	1	3
No				Yes

Did the possiblity to unlock summary of responses improve your engagement and motivation?

0	1	2	0	3
No				Yes

Did the possibility to ask your own questions improve your engagement and motivation?

0	2	1	1	2
No				Yes

Table 7. Evaluation of the gamification elements

Appendix D. Results of study at GamifIR 2015

Registration

Are you happy to be at GamifIR?

0	0	0	1	4
Unhappy				Happy

Are you sleepy?

0	3	0	0	0
Sleepy				Full of energy

Has the registration been ok?

0	0	0	3	1
Not so good				Perfect

Listening to a presentation

Have you been focused on the presentation?

1	2	0	9	10
Distracted				Focused

Is the presentation inspiring you to further research?

0	2	5	7	7
No				Yes

Has the presentation been interesting?

0	0	6	7	9
A little				Very interesting

Just asked a question

Has the reply been satisfying?

0	1	0	0	6
No				Yes

Did you ask the question because of...

1	1	2	0	3
Pure curiosity				Professional interest

Coffee break

Are you having a good time?

0	0	2	1	4
Bad				Great

Have you met people that you would like to work with?

0	2	2	1	0
Nobody				Many

Discussion

Has the discussion been interesting?

0	0	0	4	2
Boring				Interesting

Do you feel inspired?

0	1	1	3	1
Not at all				Very much

Lunch

Are you enjoying talking to people during the lunch?

0	0	0	2	4
No				Yes

Has the lunch been good?

0	1	0	3	1
Awful				Delicious

Are you happy to be at GamifIR so far?

0	0	0	1	5
Unhappy				Happy

Keynote

Do you find the topic of the keynote interesting?

0	0	1	1	7
A little				Very interesting

Table . Responses to study at GamifIR

Appendix E. Database details

Studies

This table contains the list of studies. Studies are assigned to a researcher (user) and consist of one or more question groups (events).

Table overview

Column	Type
id	int(11)
Name	varchar(255)
Description	text
ResearcherID	int(11)
SubscriptionCode	int(11)
Status	varchar(11)
question_group_count	int(11)
member_count	int(11)

SQL create query

```
CREATE TABLE IF NOT EXISTS `studies` (
  `id` int(11) NOT NULL AUTO_INCREMENT,
  `Name` varchar(255) NOT NULL,
  `Description` text,
  `ResearcherID` int(11) unsigned NOT
NULL,
  `SubscriptionCode` int(11) NOT NULL,
  `Status` int(11) NOT NULL DEFAULT
'1',
  `question_group_count` int(11) NOT
NULL,
  `member_count` int(11) NOT NULL,
  PRIMARY KEY (`id`),
  KEY `ResearcherID` (`ResearcherID`))
```

Question_groups (Events)

This table contains the list of events. Each event is assigned to a study and contains one or more questions.

Table overview

Column	Type
id	int(11)
Name	varchar(255)
StudyID	Int(11)
ResearcherID	int(11)
question_count	int(11)
member_count	int(11)

SQL create query

```
CREATE TABLE IF NOT EXISTS
`question_groups` (
  `id` int(11) NOT NULL AUTO_INCREMENT,
  `Name` varchar(255) NOT NULL DEFAULT
'Unspecified',
  `StudyID` int(11) NOT NULL DEFAULT
'1',
  `ResearcherID` int(11) NOT NULL,
  `question_count` int(11) NOT NULL,
  PRIMARY KEY (`id`),
  KEY `StudyID` (`StudyID`))
```

Questions

This table contains the list of questions. Questions are assigned to events.

Table overview

Column	Type
id	int(11)
GroupID	int(11)
ResearcherID	int(11)

SQL create query

```
CREATE TABLE IF NOT EXISTS `questions`
(
  `id` int(11) NOT NULL AUTO_INCREMENT,
  `GroupID` int(11) NOT NULL,
  `ResearcherID` int(11) NOT NULL,
```

Content	varchar(255)
Comment	text
Positive	varchar(32)
Negative	varchar(32)
answer_count	varchar(11)

```
`Content` varchar(255) NOT NULL,
`Comment` text,
`Positive` varchar(32) DEFAULT 'Yes',
`Negative` varchar(32) DEFAULT 'No',
`answer_count` int(11) NOT NULL,
PRIMARY KEY (`id`),
KEY `GroupID` (`GroupID`))
```

Statistics

I have created a separate table for the responses summaries. Each row in this table corresponds to a row in the questions table. This separation allows storage of precalculated values of total responses so that they do not need to be recalculated each time. The calculation is done by one of the CakePHP methods and can be triggered on demand, also from the web services.

Table overview

Column	Type
id	int(11)
Count1	varchar(255)
Count2	text
Count3	int(11)
Count4	int(11)
Count5	varchar(11)

SQL create query

```
CREATE TABLE IF NOT EXISTS `statistics` (
  `id` int(11) NOT NULL,
  `Count1` int(11) NOT NULL DEFAULT '0',
  `Count2` int(11) NOT NULL DEFAULT '0',
  `Count3` int(11) NOT NULL DEFAULT '0',
  `Count4` int(11) NOT NULL DEFAULT '0',
  `Count5` int(11) NOT NULL DEFAULT '0',
  PRIMARY KEY (`id`))
```

Answers

This is a simple table that stores the responses to all questions. This table is supposed to contain the biggest number of rows.

Table overview

Column	Type
id	int(11)
Value	int(11)
MemberID	int(11)
timestamp	timestamp

SQL create query

```
CREATE TABLE IF NOT EXISTS `answers` (
  `id` int(11) NOT NULL AUTO_INCREMENT,
  `Value` int(11) DEFAULT NULL,
  `QuestionID` int(11) NOT NULL,
  `MemberID` int(11) NOT NULL,
  `timestamp` timestamp NOT NULL
DEFAULT CURRENT_TIMESTAMP,
  PRIMARY KEY (`id`),
  KEY `QuestionID` (`QuestionID`),
  KEY `MemberID` (`MemberID`))
```

Users

This table contains the information about the users of the web panel. Both administrators and researchers are stored in the same table. The system recognizes their role based on the role column value when they log in.

Table overview

Column	Type
id	int(11)
username	varchar(50)
password	varchar(255)
Email	varchar(255)
Name	varchar(255)
role	varchar(20)
status	int(1)
InstitutionID	int(11)
CountryID	int(11)
study_count	int(11)

SQL create query

```
CREATE TABLE IF NOT EXISTS `users`
(
  `id` int(11) unsigned NOT NULL
AUTO_INCREMENT,
  `username` varchar(50) DEFAULT
NULL,
  `password` varchar(255) DEFAULT
NULL,
  `Email` varchar(255) NOT NULL,
  `Name` varchar(255) NOT NULL,
  `role` varchar(20) DEFAULT NULL,
  `status` int(1) NOT NULL DEFAULT
'0',
  `InstitutionID` int(11) DEFAULT
NULL,
  `CountryID` int(11) NOT NULL,
  `study_count` int(11) NOT NULL,
  PRIMARY KEY (`id`),
  KEY `CountryID` (`CountryID`))
```

Members

While the users table stores users of the web part, the members table contains the information about the mobile app users. It used to be more complex, but utilizing Facebook API for login made a couple of columns no longer needed.

Table overview

Column	Type
id	int(11)
Username	varchar(255)
Email	varchar(255)
Status	int(11)
study_count	int(11)
answer_count	int(11)

SQL create query

```
CREATE TABLE IF NOT EXISTS `members` (
  `id` int(11) NOT NULL,
  `Username` varchar(255) NOT NULL,
  `Email` varchar(255) NOT NULL DEFAULT
'NULL',
  `Status` int(11) NOT NULL DEFAULT
'1',
  `study_count` int(11) DEFAULT NULL,
  `answer_count` int(11) DEFAULT NULL,
  PRIMARY KEY (`id`))
```

Members_studies

The purpose of this table is to support the ManyToMany relation between the Members and Studies (a member subscribes to many studies and many members subscribe to a study). CakePHP automatically handles the relation if such table existists.

Table overview

Column	Type
id	int(11)
member_id	int(11)
study_id	int(11)

SQL create query

```
CREATE TABLE IF NOT EXISTS
`members_studies` (
  `ID` int(11) NOT NULL AUTO_INCREMENT,
  `member_id` int(11) NOT NULL,
  `study_id` int(11) NOT NULL,
  PRIMARY KEY (`ID`))
```

Institutions

This is an auxiliary table for the users table. Each researcher is assigned to an institution – a university, an organization, a company etc.

Table overview

Column	Type
id	int(11)
Name	varchar(255)
CountryID	int(11)
user_count	int(11)

SQL create query

```
CREATE TABLE IF NOT EXISTS
`institutions` (
  `ID` int(11) NOT NULL
AUTO_INCREMENT,
  `Name` varchar(255) NOT NULL,
  `CountryID` int(11) NOT NULL,
  `user_count` int(11) NOT NULL,
  PRIMARY KEY (`ID`),
  KEY `institutions_ibfk_1`
(`CountryID`))
```

Countries

This is an auxiliary table for the users and institutions tables.

Table overview

Column	Type
id	int(11)
country_code	varchar(2)
country_name	varchar(100)
user_count	int(11)
institution_count	int(11)

SQL create query

```
CREATE TABLE IF NOT EXISTS `countries`
(
  `id` int(11) NOT NULL
AUTO_INCREMENT,
  `country_code` varchar(2) NOT NULL
DEFAULT '',
  `country_name` varchar(100) NOT NULL
DEFAULT '',
  `user_count` int(11) NOT NULL,
  `institution_count` int(11) NOT
NULL,
  PRIMARY KEY (`id`))
```

References

1. Körber, M. and K. Bengler. *Measurement of momentary user experience in an automotive context*. in *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2013. ACM.
2. Alalawi, M.S., *User Experience: Summary Judgments and Momentary Emotions.* SPARC 2007, 2007: p. 1.
3. Conner, T.S. and L.F. Barrett, *Trends in ambulatory self-report: the role of momentary experience in psychosomatic medicine.* Psychosomatic medicine, 2012. 74(4): p. 327.
4. Hayes, G.R., et al., *Estrellita: a mobile capture and access tool for the support of preterm infants and their caregivers.* ACM Transactions on Computer-Human Interaction (TOCHI), 2014. 21(3): p. 19.
5. Csikszentmihalyi, M. and R. Larson, *Validity and reliability of the experience-sampling method.* The Journal of nervous and mental disease, 1987. 175(9): p. 526-536.
6. Karahasanović, A., et al., *Co-creation and user-generated content–elderly people's user requirements.* Computers in Human Behavior, 2009. 25(3): p. 655-678.
7. Keele, S., *Guidelines for performing Systematic Literature Reviews in Software Engineering*. 2007, EBSE Technical Report EBSE-2007-01.
8. Csikszentmihalyi, M., *Flow*. 2014: Springer.
9. Vermeeren, A.P.O.S., et al., *User experience evaluation methods: current state and development needs*, in *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*. 2010, ACM: Reykjavik, Iceland. p. 521-530.
10. Deng, L., et al., *User experience, satisfaction, and continual usage intention of IT.* European Journal of Information Systems, 2010. 19(1): p. 60-75.
11. Deterding, S., et al. *From game design elements to gamefulness: defining gamification*. in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*. 2011. ACM.
12. Machnik, M., M. Riegler, and S. Sen. *Crowdpinion: Motivating People to Share Their Momentary Opinion*. in *Second International Workshop on Gamification for Information Retrieval (GamifIR'15)*. 2015. CEUR Workshop Proceedings.
13. Graham, D., et al., *Foundations of software testing: ISTQB certification*. 2008: Course Technology Cengage Learning.
14. Bevan, N. *Classifying and selecting UX and usability measures*. in *International Workshop on Meaningful Measures: Valid Useful User Experience Measurement*. 2008.
15. Larson, R. and M. Csikszentmihalyi, *The experience sampling method.* New Directions for Methodology of Social & Behavioral Science, 1983.
16. Hektner, J.M., J.A. Schmidt, and M. Csikszentmihalyi, *Experience sampling method: Measuring the quality of everyday life*. 2007: Sage.
17. Scollon, C.N., C.-K. Prieto, and E. Diener, *Experience sampling: promises and pitfalls, strength and weaknesses*, in *Assessing well-being*. 2009, Springer. p. 157-180.
18. Csikszentmihalyi, M., S. Abuhamdeh, and J. Nakamura, *Flow*, in *Flow and the Foundations of Positive Psychology*. 2014, Springer Netherlands. p. 227-238.
19. Walsh, T., P. Nurkka, and R. Walsh. *Cultural differences in smartphone user experience evaluation*. in *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia*. 2010. ACM.
20. Barnum, C.M. and L.A. Palmer, *Tapping into Desirability in User Experience.* Usability of Complex Information Systems, 2010: p. 253.

21. Meschtscherjakov, A., A. Weiss, and T. Scherndl, *Utilizing Emoticons on mobile devices within ESM studies to measure emotions in the field.* Proc. MME in conjunction with MobileHCI, 2009. 9.
22. Obrist, M., A. Meschtscherjakov, and M. Tscheligi, *User experience evaluation in the mobile context*, in *Mobile TV: Customizing Content and Experience*. 2010, Springer. p. 195-204.
23. Estellés-Arolas, E. and F. González-Ladrón-de-Guevara, *Towards an integrated crowdsourcing definition.* Journal of Information science, 2012. 38(2): p. 189-200.
24. Howe, J., *The rise of crowdsourcing.* Wired magazine, 2006. 14(6): p. 1-4.
25. Kleemann, F., G.G. Voß, and K. Rieder, *Un (der) paid innovators.* 2008.
26. *Amazon Mechanical Turk*. 2015 [cited 2015]; Available from: https://www.mturk.com.
27. Heer, J. and M. Bostock. *Crowdsourcing graphical perception: using mechanical turk to assess visualization design*. in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2010. ACM.
28. Kittur, A., E.H. Chi, and B. Suh. *Crowdsourcing user studies with Mechanical Turk*. in *Proceedings of the SIGCHI conference on human factors in computing systems*. 2008. ACM.
29. Grady, C. and M. Lease. *Crowdsourcing document relevance assessment with mechanical turk*. in *Proceedings of the NAACL HLT 2010 workshop on creating speech and language data with Amazon's mechanical turk*. 2010. Association for Computational Linguistics.
30. Shank, D.B., *Using Crowdsourcing Websites for Sociological Research: The Case of Amazon Mechanical Turk.* The American Sociologist, 2015: p. 1-9.
31. Brabham, D.C., *Moving the crowd at iStockphoto: The composition of the crowd and motivations for participation in a crowdsourcing application.* First monday, 2008. 13(6).
32. Brabham, D.C., *Moving the crowd at Threadless: Motivations for participation in a crowdsourcing application.* Information, Communication & Society, 2010. 13(8): p. 1122-1145.
33. Poetz, M.K. and M. Schreier, *The value of crowdsourcing: can users really compete with professionals in generating new product ideas?* Journal of Product Innovation Management, 2012. 29(2): p. 245-256.
34. Eickhoff, C. and A. de Vries, *How crowdsourcable is your task.* Proceedings of the workshop on crowdsourcing for search and data mining (CSDM) at the fourth ACM international conference on web search and data mining (WSDM), 2011: p. 11-14.
35. Wang, G., et al. *Serf and turf: crowdturfing for fun and profit*. in *Proceedings of the 21st international conference on World Wide Web*. 2012. ACM.
36. Surowiecki, J., *The wisdom of crowds*. 2005: Anchor.
37. Bihr, P., *Crowdsourcing 101: Crowdsourcing vs Wisdom of the Crowd.* Slides available online at http://www. slideshare. net/thewavingcat/101-crowdsourcing-vswisdom-of-the-crowd (accessed on: 27/04/2015), 2010.
38. Zyda, M., *From visual simulation to virtual reality to games.* Computer, 2005. 38(9): p. 25-32.
39. Raddick, M.J., et al., *Galaxy zoo: Exploring the motivations of citizen science volunteers.* Astronomy Education Review, 2010. 9(1): p. 010103.
40. Cooper, S., et al., *The challenge of designing scientific discovery games.* computing, 2010. 9: p. 16.
41. Lux, M., M. Guggenberger, and M. Riegler. *Picturesort: gamification of image ranking*. in *Proceedings of the First International Workshop on Gamification for Information Retrieval*. 2014. ACM.
42. Riegler, M., et al., *Mobile Picture Guess: A Crowdsourced Serious Game for Simulating Human Perception*, in *Social Informatics*, L.M. Aiello and D. McFarland, Editors. 2015, Springer International Publishing. p. 461-468.
43. Schreiner, K., *Digital games target social change.* Computer Graphics and Applications, IEEE, 2008. 28(1): p. 12-17.
44. Ferri, G. and R. Fusaroli. *Which Narrations for Persuasive Technologies? Habits and Procedures in Ayiti: The Cost of Life*. in *AAAI Spring Symposium: Intelligent Narrative Technologies II*. 2009.

45. Burak, A., E. Keylor, and T. Sweeney, *PeaceMaker: A video game to teach peace*, in *Intelligent technologies for interactive entertainment*. 2005, Springer. p. 307-310.
46. Games, P. *The Arcade Wire: Oil God*. Available from: http://www.persuasivegames.com/games/game.aspx?game=arcadewireoil.
47. Moazzam, W., et al., *Scientific Hangman: Gamifying Scientific Evidence for General Public.* 2015.
48. Kumar, J., *Gamification at work: designing engaging business software*. 2013: Springer.
49. Antin, J. and E.F. Churchill. *Badges in social media: A social psychological perspective*. in *CHI 2011 Gamification Workshop Proceedings (Vancouver, BC, Canada, 2011)*. 2011.
50. Eickhoff, C., et al., *Quality through Flow and Immersion: Gamifying Crowdsourced Relevance Assessments.* 2012.
51. He, J., et al. *Studying user browsing behavior through gamified search tasks*. in *Proceedings of the First International Workshop on Gamification for Information Retrieval*. 2014. ACM.
52. Shovman, M., *The game of search: what is the fun in that?*, in *Proceedings of the First International Workshop on Gamification for Information Retrieval*. 2014, ACM: Amsterdam, The Netherlands. p. 46-48.
53. Radu, A.L., et al., *High Quality Photo Collection via Gamification.* 2015.
54. Kamasheva, A., et al., *Usage of Gamification Theory for Increase Motivation of Employees.* Mediterranean Journal of Social Sciences, 2015. 6(1S3): p. 77.
55. Kim, S., *Recent Advances in Gamification Application.* Advances in Information Sciences & Service Sciences, 2013. 5(13).
56. Eickhoff, C., *Human Intelligence in Search and Retrieval.*
57. Sutcliffe, A. and J. Hart, *Some Reflections on Evaluating Users' Experience.* which is commonly known as TwinTide (Towards the Integration of Trans-sectorial IT Design: p. 67.
58. Diamantaki, K., et al. *Integrating game elements for increasing engagement and enhancing User Experience in a smart city context1*. in *Workshop Proceedings of the 9th International Conference on Intelligent Environments*. 2013. IOS Press.
59. Consolvo, S. and M. Walker, *Using the experience sampling method to evaluate ubicomp applications.* IEEE Pervasive Computing, 2003. 2(2): p. 24-31.
60. Boren, T. and J. Ramey, *Thinking aloud: Reconciling theory and practice.* Professional Communication, IEEE Transactions on, 2000. 43(3): p. 261-278.
61. Sauka, F., *UX Design-How to design for and evaluate experiences?* User Behavior: p. 35.
62. Koca, A., et al. *Grasping product pragmatics: a case with internet on tv*. in *Proceedings of the 1st international conference on Designing interactive user experiences for TV and video*. 2008. ACM.
63. Schulze, K. and H. Krömker. *A framework to measure user experience of interactive online products*. in *Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research*. 2010. ACM.
64. Fierley, R. and S. Engl. *User experience methods and games: Lessons learned*. in *Proceedings of the 24th BCS Interaction Specialist Group Conference*. 2010. British Computer Society.
65. Vermeeren, A., et al. *Comparing UX measurements, a case study*. in *Proceedings of the International Workshop on Meaningful Measures: Valid Useful Experience Measurement, Reykjavik, Iceland, June*. 2008. Citeseer.
66. Dhir, A. and M. Al-kahtani, *A Case Study on User Experience (UX) Evaluation of Mobile Augmented Reality Prototypes.* Journal of Universal Computer Science, 2013. 19(8): p. 1175-1196.
67. Arhippainen, L., M. Pakanen, and S. Hickey, *Mixed UX Methods Can Help to Achieve Triumphs.* which is commonly known as TwinTide (Towards the Integration of Trans-sectorial IT Design: p. 83.
68. Idyawati, H., M. Mahmud, and A. Tap. *User Experience: Assessing the effectiveness of Internet booking service*. in *User Science and Engineering (i-USEr), 2010 International Conference on*. 2010. IEEE.

69. Carabalona, R., et al., *Light on! Real world evaluation of a P300-based brain–computer interface (BCI) for environment control in a smart home.* Ergonomics, 2012. 55(5): p. 552-563.
70. Bobkowska, A. *On explaining intuitiveness of software engineering techniques with user experience concepts*. in *Proceedings of the International Conference on Multimedia, Interaction, Design and Innovation*. 2013. ACM.
71. Strasser, E., et al., *Combining Implicit and Explicit Methods for the Evaluation of an Ambient Persuasive Factory Display*, in *Ambient Intelligence*. 2012, Springer. p. 113-128.
72. Wu, W., et al. *Quality of experience in distributed interactive multimedia environments: toward a theoretical framework*. in *Proceedings of the 17th ACM international conference on Multimedia*. 2009. ACM.
73. Keskinen, T., et al. *Evaluating the experiential user experience of public display applications in the wild*. in *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia*. 2013. ACM.
74. Abeele, V.V., B. Zaman, and D. De Grooff, *User eXperience Laddering with preschoolers: unveiling attributes and benefits of cuddly toy interfaces.* Personal and Ubiquitous Computing, 2012. 16(4): p. 451-465.
75. Harjumaa, M. and M. Isomursu, *Field Work With Older Users-Challenges in Design and Evaluation of Information Systems.* Electronic Journal of Information Systems Evaluation, 2012. 15(1).
76. Bardzell, J., et al. *Making user engagement visible: a multimodal strategy for interactive media experience research*. in *CHI'08 Extended Abstracts on Human Factors in Computing Systems*. 2008. ACM.
77. Springett, M. *Assessing user experiences within interaction: experience as a qualitative state and experience as a causal event*. in *International Workshop on*. 2008. Citeseer.
78. Goh, K.N., et al., *Designing a Checklist for an E-Commerce Website Using Kansei Engineering*, in *Advances in Visual Informatics*. 2013, Springer. p. 483-496.
79. Walsh, T. and P. Nurkka. *Approaches to cross-cultural design: two case studies with UX web-surveys*. in *Proceedings of the 24th Australian Computer-Human Interaction Conference*. 2012. ACM.
80. Hu, J., M. Janse, and H.-j. Kong. *User experience evaluation of a distributed interactive movie*. in *HCI International*. 2005.
81. van de Laar, B., et al., *Brain–Computer Interfaces and User Experience Evaluation*, in *Towards Practical Brain-Computer Interfaces*. 2013, Springer. p. 223-237.
82. Väätäjä, H., T. Koponen, and V. Roto. *Developing practical tools for user experience evaluation: a case from mobile news journalism*. in *European Conference on Cognitive Ergonomics: Designing beyond the Product---Understanding Activity and User Experience in Ubiquitous Environments*. 2009. VTT Technical Research Centre of Finland.
83. Korhonen, H., J. Arrasvuori, and K. Väänänen-Vainio-Mattila. *Analysing user experience of personal mobile products through contextual factors*. in *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia*. 2010. ACM.
84. Khan, K. *User Experience in Mobile Phones by Using Semantic Differential Methodology*. in *6th European Conference on Information Management and Evaluation*. 2012. Academic Conferences Limited.
85. Park, J., et al., *Developing elements of user experience for mobile phones and services: survey, interview, and observation approaches.* Human Factors and Ergonomics in Manufacturing & Service Industries, 2011.
86. Kujala, S., et al., *UX Curve: A method for evaluating long-term user experience.* Interacting with Computers, 2011. 23(5): p. 473-483.
87. Wakkary, R. and M. Hatala. *ec (h) o: Situated Play in a Tangible and Audio Museum Guide*. in *Proceedings of the 6th Conference on Designing Interactive systems*. 2006. ACM.

88. Olsson, T. and M. Salo. *Narratives of satisfying and unsatisfying experiences of current mobile augmented reality applications*. in *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*. 2012. ACM.
89. Ariely, D. and Z. Carmon, *The sum reflects only some of its parts: A critical overview of research about summary assessment of experiences.* Time and decision: Economic and psychological perspectives on inter-temporal choice, 2003: p. 323-350.
90. *CakePHP Cookbook 3.x documentation*. 2015 [cited 2015; Available from: http://book.cakephp.org.
91. Reenskaug, T.M.H., *The original MVC reports.* 1979.
92. *Ionic: Advanced HTML5 Hybrid Mobile App Framework*. 2015 [cited 2015; Available from: http://ionicframework.com.
93. *AngularJS - Superheroic JavaScript MVW Framework*. [cited 2015; Available from: https://angularjs.org.
94. *MySQL :: The world's most popular open source database*. 2015 [cited 2015; Available from: https://www.mysql.com.
95. Sillitti, A. and G. Succi, *Requirements engineering for agile methods*, in *Engineering and Managing Software Requirements*. 2005, Springer. p. 309-326.
96. *Apache Cordova™*. [cited 2015 17.04]; Available from: https://cordova.apache.org.
97. Coenraets, C. *OpenFB*. [cited 2015; Available from: https://github.com/ccoenraets/OpenFB.
98. Bristowe, J. *What is a Hybrid Mobile App?* 2015 [cited 2015; Available from: http://developer.telerik.com/featured/what-is-a-hybrid-mobile-app/.
99. *jQuery*. 2015 [cited 2015; Available from: https://jquery.com.
100. *JSON*. [cited 2015; Available from: http://www.json.org.
101. Behera, S. *Demo of a simple CRUD RESTful php service used with AngularJS and MySQL*. 2014 [cited 2015; Available from: http://www.angularcode.com/demo-of-a-simple-crud-restful-php-service-used-with-angularjs-and-mysql/.
102. *Representational state transfer*. 2015 [cited 2015; Available from: http://en.wikipedia.org/wiki/Representational_state_transfer.
103. *Apache™ Subversion®, "Enterprise-class centralized version control for the masses"*. 2015 [cited 2015 16.04]; Available from: https://subversion.apache.org.
104. Intille, S.S., et al. *A context-aware experience sampling tool*. in *CHI'03 extended abstracts on Human factors in computing systems*. 2003. ACM.
105. Boim, R., et al. *Asking the right questions in crowd data sourcing*. in *Data Engineering (ICDE), 2012 IEEE 28th International Conference on*. 2012. IEEE.

Table of figures